



SIANN: Strain Identification by Alignment to Near Neighbors

Samuel Minot, Stephen D Turner, Krista L Ternus, et al.

bioRxiv first posted online January 10, 2014

Access the most recent version at doi: <http://dx.doi.org/10.1101/001727>

**Creative
Commons
License**

The copyright holder for this preprint is the author/funder. It is made available under a [CC-BY-NC-ND 4.0 International license](#).

SIANN: Strain Identification by Alignment to Near Neighbors

Samuel S. Minot, Stephen D. Turner, Krista L. Ternus, and Dana R. Kadavy

January 9, 2014

Abstract

Next-generation sequencing is increasingly being used to study samples composed of mixtures of organisms, such as in clinical applications where the presence of a pathogen at very low abundance may be highly important. We present an analytical method (SIANN: Strain Identification by Alignment to Near Neighbors) specifically designed to rapidly detect a set of target organisms in mixed samples that achieves a high degree of species- and strain-specificity by aligning short sequence reads to the genomes of near neighbor organisms, as well as that of the target. Empirical benchmarking alongside the current state-of-the-art methods shows an extremely high Positive Predictive Value, even at very low abundances of the target organism in a mixed sample. SIANN is available as an Illumina BaseSpace app, as well as through Signature Science, LLC. SIANN results are presented in a streamlined report designed to be comprehensible to the non-specialist user, providing a powerful tool for rapid species detection in a mixed sample. By focusing on a set of (customizable) target organisms and their near neighbors, SIANN can operate quickly and with low computational requirements while delivering highly accurate results.

Introduction

There are many different methods that characterize the mixture of organisms present within a metagenomic dataset. Such datasets are generated when a complex environmental sample is processed by a “next-generation” high-throughput genome sequencing protocol, and they consist of large numbers of short nucleotide sequences. Each sequence represents a small fragment of a randomly selected genome from the very large collection of genomes present in the source sample. Those sequences indicate the presence of one organism or another according to their similarity to a set of known reference genomes. While a given sequence may be unique to one species, it also may be found in diverse organisms across the tree of life. Therefore, one analytical challenge (among many) is to take that collection of sequences (likely numbering in the millions) and accurately determine what species are present in the sample. Here we describe a novel method (SIANN: Strain Identification by Alignment to Near Neighbors) that is specifically designed to rapidly detect a set of targeted organisms from a metagenomic dataset by aligning reads to genomic regions that are unique at the strain or species level.

SIANN: Strain Identification by Alignment to Near Neighbors

The analytical question motivating a particular piece of metagenomic bioinformatic analysis may vary widely by user and sample type (Segata, et al., 2013). For example, the function of the human gut microbiome may depend on the relative abundance of hundreds of species of bacteria and the types of metabolic genes they contain (Wu, et al., 2011; Schloissnig, et al., 2013). In contrast, the clinical treatment of a patient may depend on whether or not a particular virus, or a consortium of co-infecting pathogens, is/are detected in their blood. It is this second class of presence/absence questions that SIANN is designed to address. SIANN is appropriate for situations in which a user wants to know whether a particular organism or set of organisms is present in a sample, but isn't interested in the functions encoded in their genomes, the relative abundance of each organism, or any other more in-depth analysis.

Methods

Approach

Metagenomic classification methods are based on a wide variety of theoretical underpinnings. The basic varieties include alignment of reads to various nucleotide databases or exact matching to nucleotide or protein signature sequences (or *kmers*). A representative set of recent methods are described in Table 1 (also see Bazinet & Cummings 2012).

| Name | Method | Reference |
|-------------------|--|--|
| MEGAN | Alignment to large nucleotide database | Huson, et al., 2011 |
| PhymmBL | Alignment to large nucleotide database with interpolated Markov models | Brady & Salzberg, 2011 |
| Metaphyler | Alignment to clade-specific marker genes | Liu, et al., 2011 |
| MetaPhlAn | Alignment to clade-specific marker genes | Segata, et al., 2012 |
| LMAT | Nucleotide kmer matching | Ames, et al., 2013 |
| Kraken | Nucleotide kmer matching | Wood & Salzberg, <i>in submission</i> |
| Sequedex | Protein kmer matching | Berendzen, et al., 2012 |
| mOTU | Alignment to universal marker genes | Sunagawa, et al., 2013 |
| Phylosift | Insertion into reference nucleotide and protein alignments | Darling, et al., <i>in preparation</i> |

Table 1. Summary of methods for metagenomic classification.

Overall, these methods are designed to either classify individual reads to, and/or predict the total abundance of, clades (e.g. genus or species) across the entire tree of life. They generally require reference databases that are very large and/or require a large amount of processing to generate. The gap SIANN is designed to fill is when the entire tree of life is irrelevant, and only predefined subsets of organisms need to be detected. For an underlying method we chose read alignment to diagnostic genomic regions because the algorithms for read alignment are highly parallelizable and have been optimized heavily by the community at large (the current implementation of

SIANN: Strain Identification by Alignment to Near Neighbors

SIANN uses bowtie2 [Langmead & Salzberg, 2012] for the alignment function, but can be adapted to any alignment algorithm). This approach is distinct from using clade-specific marker genes (Segata, et al., 2012) because unique regions that are larger, smaller, or outside of genes can also be used. Furthermore, this approach supports the rapid construction of custom databases using reference genome sets that require only minimal user-supplied structure.

To understand the principle at work, consider a set of reads that have been aligned to the genomes of several strains belonging to two species. Some regions of those genomes are species-specific, some are strain-specific, and some are shared (Figure 1a). When a set of reads is aligned to those genomes such that each read is placed in as many locations as it has a match (at a reasonably stringent threshold), visual inspection of the distribution of reads yields an intuitive understanding of the true source organism as Species I/Strain B (Figure 1b). If Strain B were not present in the reference database, it would still be clear that the organism was an unknown strain of Species I.

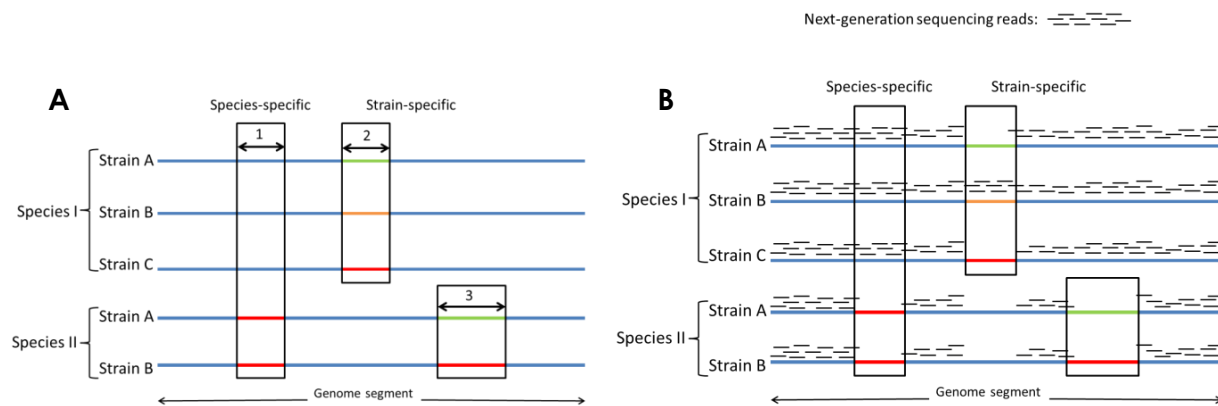


Figure 1. A) For a group of strains belonging to two different species, some regions may be unique to each species (region 1), while other regions may be unique to strains within each species (regions 2 and 3). B) A set of reads are aligned to these genomes, and the ones that align in a species- or strain-specific manner are identified by the combination of genomes to which they align. In this example, Strain B of Species I is the organism identified.

The unique identification of a species or strain is quantified by the proportion of the genome that is determined to be species- or strain-specific (defined as reads that are aligned to regions that are species- or strain-specific). Each species and strain is then assigned a numerical measure of the proportion that is covered by these diagnostic reads, and that proportional measure is compared to the ideal case, where sequences from a single organism (generated *in silico*) are aligned against the database in an identical manner. After that normalization factor is applied, the resulting score indicates whether the source sample contained any of the organisms in the reference database.

SIANN: Strain Identification by Alignment to Near Neighbors

The analysis is conducted independently on both the species and the strain level, so that if the true strain is not present in the database, the species of origin will still be identified. While many methods consider the complete taxonomic tree and assign reads to the least common ancestor, SIANN considers only two taxonomic levels: species and strain, throwing out anything that is not unique at one of those levels and thus obviating many of the confounding factors introduced by manually curated taxonomies.

The example shown in Figure 1b indicates that species-specific reads are identified as reads that align to one species (Species I, in that case) but not the other. If Species II were not present in the example shown in Figure 1b, a much larger number of reads would be assigned as "species-specific," when in fact those regions are shared with other species. Therefore, the ability of this method to identify strain- and species-specific sequences is a direct function of the inclusion of near neighbors in the reference database. This characteristic is shared among many classification algorithms, but it is of particular note for this method when users have an opportunity to construct their own database.. In order to detect a target species with a high degree of specificity (reducing false positives), it is necessary to include other related species in the reference database. Only by parallel alignment to those near neighbors can the redundant sequences be separated from the species-specific ones. For example, in order to detect *Bacillus anthracis* in a sample, it would be necessary to include other species of *Bacilli* in the reference database so that the presence of *B. cereus* or *B. thuringiensis* in a sample does not lead to a false call for *B. anthracis*.

The nomenclature of genus, species, and strain is potentially problematic because it does not correspond to a consistent degree of evolutionary distance or genomic distinctiveness. The ability to distinguish two organisms by any method using genomic sequence data is proportional to the amount of each genome that is shared or unique. One might assume that any two organisms of the same species will have a relatively predictable amount of shared genomic identity. However, some pairs of organisms from the same species may have less in common than other pairs of organisms from different species or even genera. This ambiguity impacts SIANN in two ways. If two organisms have very little genomic sequence to distinguish them, the sensitivity of SIANN to detect either one will diminish (the rate of false negatives will increase as the likelihood of sequencing unique regions decreases). Conversely, if an organism is extremely dissimilar to the near neighbors selected for the database, the specificity with which SIANN detects that organism will decline (the rate of false positives will increase as the number of related genomes available in the database decreases). For example, if a database contained only *E. coli* and *B. anthracis*, a sample containing *B. cereus* would be misidentified as containing *B. anthracis*. In the intended use case, a database targeting *B. anthracis* would contain *B. cereus* and a number of other near neighbors to prevent that kind of misidentification. It would be convenient to say that

SIANN: Strain Identification by Alignment to Near Neighbors

an ideal database can be made by calculating the ideal genetic distance between all references and then finding an ideal set of organisms to make up that database, but the behavior of any database will be governed by the particular genomes of the organisms it encounters in the wild. Because not all organisms evolve in the same manner (differences in mutation rate, horizontal gene transfer, recombination, etc), the suitability of a database and method to detect a given organism can only be determined by thorough validation and benchmarking, as well as updating the reference database as needed. Users of SIANN may construct their own custom databases to include newly identified genomes or specific subsets of genomes that best suit their research interests.

Steps to construct a custom database:

1. Select a set of target organisms
2. Gather a set of genome sequences for those target organisms as well as a matched set of near neighbors
3. Using those reference genome sequences as an input, SIANN will:
 - a. Construct a reference index for alignment
 - b. Simulate a set of reads from each genome
 - c. Align each of those simulated read sets to all of the reference genomes
 - d. Calculate the proportion of each reference genome that is strain- or species-specific
 - e. [If two organisms do not have a minimal amount of unique sequences that exceeds the rate of sequencing error, SIANN asks that all but one of those organisms are removed from the database to eliminate redundancy. Note that the user can provide a single representative genome with multiple strain names so that the redundant strain names are not lost.]

The files contained within each SIANN database are a compressed genomic index and a list containing the proportion of each reference genome that was found to be strain- or species-specific during database construction.

To run SIANN:

1. Select a pre-made SIANN database and a set of sequences to be analyzed, and
2. SIANN will:
 - a. Align each of the reads against the reference genomes
 - b. Calculate the proportion of each reference genome that is strain- or species-specific within those reads
 - c. Compare that proportion to the simulated ideal case generated during database creation

SIANN: Strain Identification by Alignment to Near Neighbors

- d. Calculate the probability that the given results could be generated by random chance
- e. Report the normalized proportion and non-parametric statistic of likelihood for each strain and species in the reference database. The normalized proportion of the genome covered by strain- or species-specific reads is the primary statistic reported by this tool.

Benchmarking

The performance of SIANN (version 1.6) was tested in comparison to the following state-of-the-art metagenomic classification programs: LMAT (version 1.2), MetaPhlAn (version 1.7.7), and Kraken (version 0.9.1b). All of the programs in Table 1 were investigated for this effort, and three were chosen based on their ability to run on our high-performance computing cluster with an execution time and memory requirement that would be suitable to a clinical lab. Each program was run on a set of 600 simulated datasets generated by MetaSim (Richter, et al., 2008). Each dataset consisted of 15,000,000 reads (100bp single-ended) with Illumina-simulated error (fourth-degree polynomial) (Korbel, et al., 2009). The 600 datasets were broken into 12 sets of 50 replicates. Each of the 12 sets contained organisms at different levels of abundance as shown in Table 2. Organisms were specifically chosen in pairs so that the ability to distinguish these near neighbors could be determined. The abundances were staggered at 4-fold intervals so that a wide range could be evaluated. All known species of near neighbors for each of

| Organism | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|--------------------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| <i>Bacillus anthracis</i> | 0.074% | 0.30% | 1.2% | 4.7% | 19% | 76% | | | | | | |
| <i>Bacillus cereus</i> | | | | | | | 0.074% | 0.30% | 1.2% | 4.7% | 19% | 76% |
| Hanta virus | | 1.2% | 4.7% | 19% | 76% | 0.074% | 0.30% | | | | | |
| Rift valley fever virus | 0.30% | | | | | | | 1.2% | 4.7% | 19% | 76% | 0.074% |
| <i>Clostridium botulinum</i> | | | 19% | 76% | 0.074% | 0.30% | 1.2% | 4.7% | | | | |
| <i>Clostridium difficile</i> | 1.2% | 4.7% | | | | | | | 19% | 76% | 0.074% | 0.30% |
| <i>Listeria fleischmannii</i> | | | | 0.074% | 0.30% | 1.2% | 4.7% | 19% | 76% | | | |
| <i>Listeria monocytogenes</i> | 4.7% | 19% | 76% | | | | | | | 0.074% | 0.30% | 1.2% |
| Monkeypox virus | | | | | 1.2% | 4.7% | 19% | 76% | 0.074% | 0.30% | | |
| <i>Vaccinia virus</i> | 19% | 76% | 0.074% | 0.30% | | | | | | | 1.2% | 4.7% |
| <i>Yersinia enterocolitica</i> | | | | | | 19% | 76% | 0.074% | 0.30% | 1.2% | 4.7% | |
| <i>Yersinia pestis</i> | 76% | 0.074% | 0.30% | 1.2% | 4.7% | | | | | | | 19% |

Table 2. The abundance of each target organism in each set of simulated datasets. Each set is indicated by the number in the top row, and was generated with 50 replicates.

SIANN: Strain Identification by Alignment to Near Neighbors

the 12 target organisms were included in the reference database used by SIANN for this benchmarking ("Target Pathogen Database") and are shown in Appendix 1.

Each program outputs a distinct measure. Kraken and LMAT both count the reads assigned to each taxon, MetaPhlAn calculates the abundance, and SIANN outputs a measure of the proportion of diagnostic genomic regions present. To put these measures on an even footing, we empirically calculated the false positive rate for each method over all 600 samples, at each possible measure of output. Because each dataset is made up of known organisms, any result can be classified as true or false. Therefore, for any possible result (say, 513 reads classified by LMAT or 1.6% abundance assigned by MetaPhlAn), one can calculate the proportion of calls with at least the same amount of support that were correct (True Positives/[True Positives+False Positives]), over all of the 600 datasets. That measure is commonly given as Positive Predictive Value (PPV). For each program, the results can be translated from the raw value into a PPV that is based on this empirical measure of error. The key item of interest is the PPV value for the results that we know to be true positives, the defined spike organisms. Another way of describing this approach is to say that the results of each program have been normalized to the false positive error rate that was empirically observed. If another set of samples were generated, the PPV vs. raw value curve (Figure 2) would likely fall differently, but in this case it gives us a means of comparing a diverse set of methods against the same ground truth. If method 1 detects an organism with a higher PPV than method 2 does, it means that method 1 has fewer false positives in the range that it reports true positives, which is the definition of utility in this setting.

For each method, PPV was calculated as a function of raw output value. Briefly, this was done by compiling the output for all 600 samples, labeling each result as false or true based on the sample set that it came from, and then calculating (at each possible value of output) what the proportion of TP/[TP+FP] was for results with at least that level of raw output. Some simplification steps were taken, such as focusing on the species-level assignments (for comparison with methods that do not perform strain assignment), and only taking the top hit for each species from each dataset. Custom R and BASH scripts were used for the data compilation and analysis.

Results

The relationship of raw output value to PPV is shown for each of the four methods in Figure 2. The point at which PPV is very close to 1 (where 95% of results are true positives) is ~41,000 reads for Kraken, ~2,800 reads for LMAT, ~38% abundance for MetaPhlAn, and 0.21 for SIANN. For SIANN this means that having 38% of the species-unique genome covered by reads resulted in the vast majority of calls being accurate.

SIANN: Strain Identification by Alignment to Near Neighbors

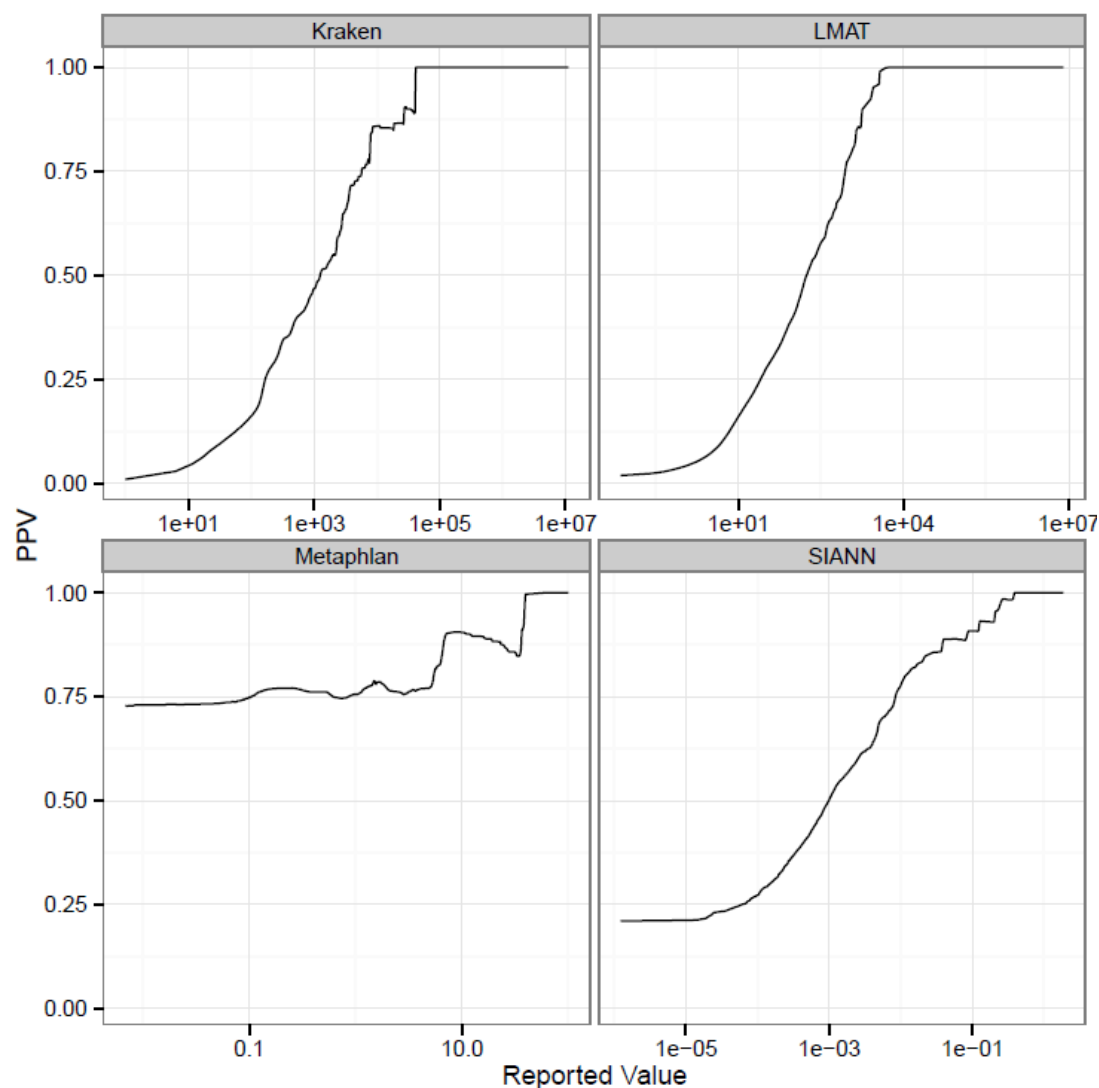


Figure 2. Relationship of reported value for each program (horizontal axis, log scale) to the empirically-determined Positive Predictive Value (PPV), shown on the vertical axis. While the exact values depend on the test data used, the general values at significant cutoff values (0.8, 0.9, 0.95 PPV) remain relatively constant across different datasets (data not shown).

SIANN: Strain Identification by Alignment to Near Neighbors

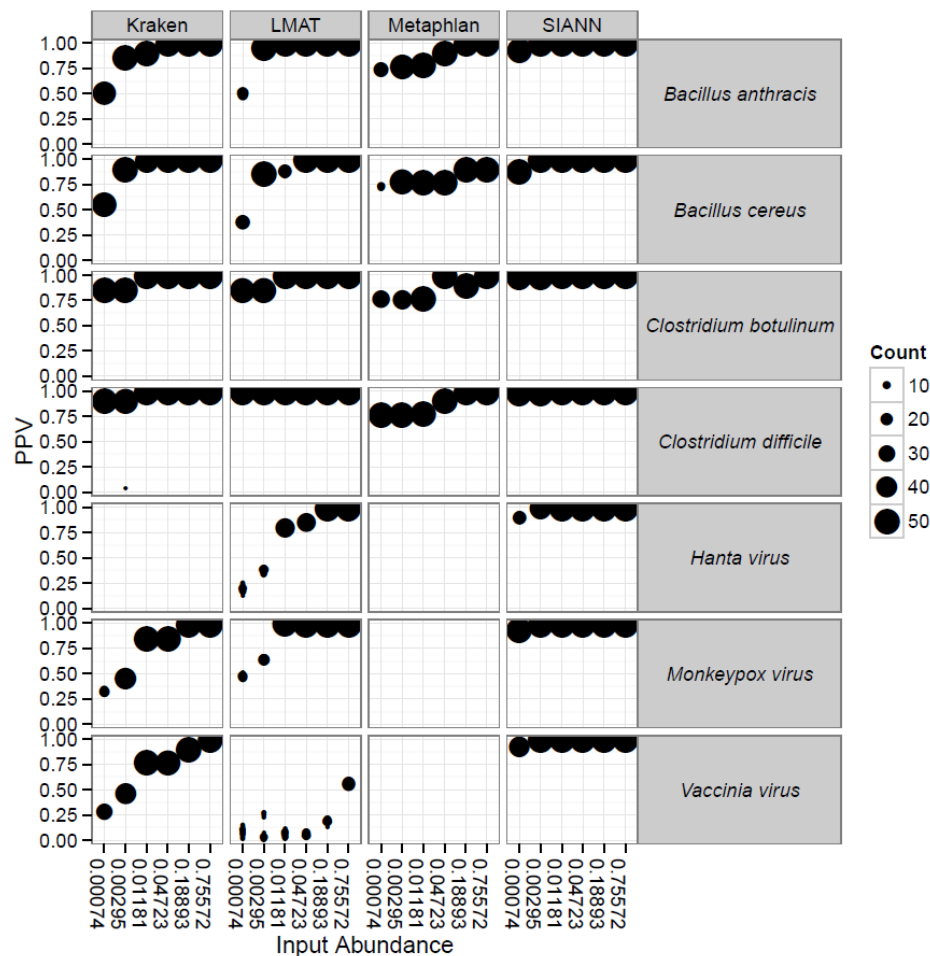


Figure 3. The Positive Predictive Value (PPV, vertical axis) is shown for each organism (boxes on right), at each level of known abundance (horizontal axis, see Table 2), and each program (boxes at top), across a maximum of 50 replicates (indicated by the size of each point). Note that the reference database for MetaPhlAn does not include viruses, and the reference database for Kraken does not include RNA viruses

extremely unlikely that a 300bp exact match would arise due to random chance, and so the user could say with confidence that the organism of interest is found within the sequence data (not considering contamination, horizontal gene transfer, etc). However, such an approach is not currently implemented in an automated method, and many of the steps needed to make that assertion are performed manually by a domain expert, including alignment to near neighbors and ensuring that the read does not fall within a transposon, plasmid, etc. Therefore, while one could say that a single read is all that is needed to state with high PPV that an organism is present, the amount of reads assigned in an automated manner needed to achieve that level of PPV will number in the thousands (Fig 2).

For read-assignment methods (such as LMAT and Kraken), manual inspection of the results may yield a different understanding of confidence than is presented here, or in any automated analysis. For example, while each read that is assigned by LMAT and Kraken fall above a certain cutoff for species-specificity, some individual reads may be much more specific than others. One could

identify a read that aligns to a single species of bacteria with 100% accuracy over its 300bp length, with the next closest match being only 90% similar. It is

SIANN: Strain Identification by Alignment to Near Neighbors

The next phase of benchmarking was to determine how many raw input reads were needed to achieve the threshold for high PPV. To demonstrate this we plotted the known abundance of each spike organism against the PPV value generated by each method (Figure 3). Each point (an organism at a known level of abundance) is comprised of a maximum of 50 replicates, where the diameter of each point increases with an increasing number of replicates. For demonstration purposes we are showing two pairs of bacteria and three viruses. Recall that for each of the pairs of bacteria (and the two poxviruses) any sample containing one did not contain the other (as shown in Table 1). The empty boxes result from the organisms not being called at any abundance. For MetaPhlAn, that is a result of no viruses being included in the version of the reference database available for this analysis. Kraken assigned no reads to Hanta virus because viral RNA genomes were not included in this version of the reference database (personal communication with D. Wood). This emphasizes the point that a) the ability to create custom databases targeting organisms of interest can be valuable, and b) the performance of any method must be benchmarked against each potential target of interest.

All methods were able identify the bulk of organisms in their databases at high abundances (75% and 18%, Figure 3), however performance varied considerably at lower abundances and depended on the particular organism and method used. SIANN detected each organism at high confidence, even at levels as low as 0.3% and 0.07% of the total.

Discussion

The process of detecting trace amounts of a specific organism in a complex mixture of DNA is challenging enough for an expert, but that pales in comparison to the difficulty of accomplishing the same certainty of detection in an automated manner. The results presented here show that SIANN rapidly detects the presence of a given set of organisms with a high degree of specificity and sensitivity. For example, at the 95% confidence (PPV) cutoff of 0.2, SIANN reliably detects all of the organisms tested here at as low as 0.3% abundance. This strong performance is likely due to the fact that SIANN is able to use a method (read alignment to whole genomes) that would be far too computationally costly if it were applied to the entire collection of known genomes. By focusing on a set of (customizable) target organisms and their near neighbors, SIANN can operate quickly and with low computational requirements while delivering highly accurate results.

SIANN is available on Illumina's BaseSpace (www.basespace.illumina.com) as a NativeApp, with the database tested here (Appendix 1), as well as a database made from the NCBI representative set of prokaryotic genomes (ftp://ftp.ncbi.nlm.nih.gov/genomes/genome_reports/) (Appendix 2) and the complete set of NCBI viral genomes (<ftp://ftp.ncbi.nlm.nih.gov/refseq/release/viral/>) (Appendix 3).

SIANN: Strain Identification by Alignment to Near Neighbors

BaseSpace was chosen as an appropriate release platform because while the entire set of software and dependencies can be deployed by the user from within a graphical user interface, the actual computation takes place in a controlled 'cloud' environment. Such a distribution strategy obviates the need to satisfy the multiple software or OS dependencies that often arises with academic computational methods. Results for SIANN are compiled into a report format, showing both the organisms that surpass 95% confidence, as well as the closest strain match for each species. The default view masks the raw data output, so that the results are human-readable and do not present extraneous information. While the code for execution and database-construction on a users system is available from Signature Science, LLC, additional databases on the BaseSpace platform can be made available upon request.

There is a neverending list of questions that one could ask of metagenomic sequencing data generated from important samples. Instead of answering them all, we demonstrate a technique with a very narrow focus that is able to report with a high degree of confidence whether a given set of organisms is present in a sample. These results are presented to the user in a comprehensible format, and accessible on a commonly-used web platform. The world of bioinformatics will continue to progress and develop more sophisticated tools for metagenomic analysis, and we hope that the utility of SIANN will convince others to package and benchmark their tools in a way that they can be used with confidence by the larger public, as well as the research community.

SIANN: Strain Identification by Alignment to Near Neighbors

References

- Ames SK, Hysom DA, Gardner SN, Lloyd GS, Gokhale MB, and Allen JE (2013). Scalable metagenomic taxonomy classification using a reference genome database. **Bioinformatics**, 29(18):2253-60. doi: 10.1093/bioinformatics/btt389.
- Bazinet AL & Cummings MP (2012). A comparative evaluation of sequence classification programs. **BMC Bioinformatics**. 13: 92. doi: 10.1186/1471-2105-13-92
- Berendzen J, Bruno WJ, Cohn JD, Hengartner NW, Kuske CR, McMahon BH, Wolinsky MA and Xie G (2012). Rapid phylogenetic and functional classification of short genomic fragments with signature peptides. **BMC Research Notes**, 5:460 doi:10.1186/1756-0500-5-460
- Brady A and Salzberg SL (2009). Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. **Nature Methods**, 6(9):673-6. doi: 10.1038/nmeth.1358.
- Huson DH, Mitra S, Weber N, Ruscheweyh H, and Schuster SC (2011). Integrative analysis of environmental sequences using MEGAN4. **Genome Research**, 21:1552-1560.
- Korbel JO, Abyzov A, Mu XJ, Carriero N, Cayting P, Zhang Z, Snyder M and Gerstein MB (2009). PEMer: a computational framework with simulation-based error models for inferring genomic structural variants from massive paired-end sequencing data. **Genome Biology**, 10:R23 doi:10.1186/gb-2009-10-2-r23
- Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. **Nature Methods**. 2012, 9:357-359.
- Liu B, Gibbons T, Ghodsi M, Treangen T, and Pop M (2011). Accurate and fast estimation of taxonomic profiles from metagenomic shotgun sequences. **BMC Genomics**, 12(Suppl 2):S4
- Richter DC, Ott F, Auch AF, Schmid R, Huson DH (2008). MetaSim: a sequencing simulator for genomics and metagenomics. **PLoS One**, 3(10):e3373. doi: 10.1371/journal.pone.0003373
- Schloissnig S, Arumugam M, Sunagawa S, Mitreva M, Tap J, Zhu A, Waller A, Mende DR, Kultima JR, Martin J, Kota K, Sunyaev SR, Weinstock GM, Bork P (2013). Genomic variation landscape of the human gut microbiome. **Nature**, 493(7430):45-50. doi: 10.1038/nature11711
- Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C (2013). Computational meta'omics for microbial community studies. **Molecular Systems Biology**, 14(9):666. doi: 10.1038/msb.2013.22

SIANN: Strain Identification by Alignment to Near Neighbors

Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, and Huttenhower C (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. **Nature Methods**, 9(8):811-4. doi: 10.1038/nmeth.2066.

Sunagawa S, Mende DR, Zeller G, Izquierdo-Carrasco F, Berger SA, Kultima JR, Coelho LP, Arumugam M, Tap J, Nielsen HB, Rasmussen S, Brunak S, Pedersen O, Guarner F, de Vos WM, Wang J, Li J, Doré J, Ehrlich SD, Stamatakis A & Bork P (2013). Metagenomic species profiling using universal phylogenetic marker genes. **Nature Methods**, 10, 1196–1199 (2013) doi:10.1038/nmeth.2693

Wood DE and Salzberg SL. Ultrafast metagenomic sequence classification using exact alignments. In submission.

Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, Sinha R, Gilroy E, Gupta K, Baldassano R, Nessel L, Li H, Bushman FD, Lewis JD (2011). Linking long-term dietary patterns with gut microbial enterotypes. **Science**. 334(6052):105-8. doi: 10.1126/science.1208344.

SIANN: Strain Identification by Alignment to Near Neighbors

Appendices

Appendix 1: Target Pathogen Database

| | | | | |
|------------------------------------|--|---|---------------------------------------|-------------------------------------|
| Arenaviridae | Arenavirus | Bacillus cytotoxicus NVH 391-98 | Bunyaviridae | Phlebovirus Rift-Valley-fever-virus |
| Golden-Gate-virus | | Bacillus mycoides DSM 2048 | Burkholderia | cenocepacia HI2424 |
| Arenaviridae | Arenavirus Lujovirus | Bacillus mycoides Rock1-4 | Burkholderia | cenocepacia J2315 |
| Arenaviridae | Flexal-virus segment-L | Bacillus thuringiensis BMB171 | Burkholderia | cenocepacia MC0-3 |
| Arenaviridae | New-world-arenaviruses Allpahuayo-virus | Bacillus thuringiensis Bt407 | Burkholderia | cepacia GG4 |
| Arenaviridae | New-world-arenaviruses Chapare-virus | Bacillus thuringiensis HD-771 | Burkholderia | gladioli BSR3 |
| Arenaviridae | New-world-arenaviruses Guanarito-virus | Bacillus thuringiensis serovar chinensis CT-43 | Burkholderia | glumae BGR1 |
| Arenaviridae | New-world-arenaviruses Junin-virus | Bacillus thuringiensis serovar konkukian 97-27 | Burkholderia | mallei 2002721280 |
| Arenaviridae | New-world-arenaviruses Machupo-virus | Brucella abortus A13334 | Burkholderia | mallei ATCC 10399 |
| Arenaviridae | New-world-arenaviruses Sabia-virus | Brucella ceti B1 94 | Burkholderia | mallei NCTC 10247 |
| Arenaviridae | New-world-arenaviruses Tacaribe-virus | Brucella ceti M13 05 1 supercont1 22 | Burkholderia | mallei SAVP1 |
| Arenaviridae | New-world-arenaviruses Whitewater-Arroyo-virus | Brucella melitensis ATCC 23457 | Burkholderia | multivorans ATCC 17616 |
| Arenaviridae | Old-world-arenaviruses lppy-virus | Brucella melitensis bv 1 str 16M | Burkholderia | oklahomensis C6786 |
| Arenaviridae | Old-world-arenaviruses Lassa-virus | Brucella ovis ATCC 25840 | Burkholderia | oklahomensis EO147 |
| Arenaviridae | Old-world-arenaviruses Mopeia-virus-AN20410 | Brucella suis 1330 | Burkholderia | pseudomallei 1026b |
| Asfarviridae | African-swine-fever-virus Benin-971-pathogenic-isolate | Brucella suis ATCC 23445 | Burkholderia | pseudomallei 1106a |
| Asfarviridae | African-Swine-Fever-Virus | Bunyaviridae Akabane-virus segment-M | Burkholderia | pseudomallei 1710b |
| Bacillus anthracis A2012 Bant 02 1 | | Bunyaviridae Hantavirus Andes-virus | Burkholderia | pseudomallei 668 |
| Bacillus anthracis Ames Ancestor | | Bunyaviridae Hantavirus Dobrava-Belgrade-virus-strain-DOBV-Ano-Poroia-Afl9-1999 | Burkholderia | pseudomallei BPC006 |
| Bacillus anthracis Sterne | | Bunyaviridae Hantavirus Hantaan-virus | Burkholderia | pseudomallei K96243 |
| Bacillus cereus 03BB102 | | Bunyaviridae Hantavirus Puumala-virus | Burkholderia | pyrrocinia CH-67 |
| Bacillus cereus AH187 | | Bunyaviridae Hantavirus Seoul-virus-strain-Seoul-80-39-clone-1 | Burkholderia | thailandensis ATCC 700388 |
| Bacillus cereus AH820 | | Bunyaviridae Hantavirus Sin-Nombre-virus | Burkholderia | thailandensis E264 |
| Bacillus cereus ATCC 14579 | | Bunyaviridae Hantavirus Thottapalayam-virus | Burkholderia | thailandensis MSMB121 |
| Bacillus cereus B4264 | | Bunyaviridae Hantavirus Tula-virus | Campylobacter coli JV20 | |
| Bacillus cereus F65185 | | Bunyaviridae Nairovirus Crimean-Congo-hemorrhagic-fever-virus | Campylobacter fetus subsp fetus 82-40 | |
| | | Bunyaviridae Nairovirus Dugbe-virus | Campylobacter jejuni RM1221 | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|--|
| Campylobacter jejuni subsp doylei 26997 | Coccidioides posadasii CPA 0020 | Coxiella burnetii Dugway 5J108 111 |
| Campylobacter jejuni subsp jejuni 81-176 | Coronaviridae | Coxiella burnetii RSA 493 |
| Campylobacter jejuni subsp jejuni NCTC 11168 ATCC 700819 | Alphacoronavirus Bat- coronavirus-HKU2 | Diplorickettsia massiliensis 20B CS 1 |
| Campylobacter upsaliensis JV21 | Coronaviridae | Ehrlichia canis str Jake |
| Clostridium acetobutylicum DSM 1731 | Alphacoronavirus Feline- infectious-peritonitis-virus | Ehrlichia chaffeensis str Arkansas |
| Clostridium botulinum A str ATCC 3502 | Coronaviridae | Ehrlichia chaffeensis str Sapulpa ctg90 |
| Clostridium botulinum BKT015925 | Alphacoronavirus Human- coronavirus-229E | Ehrlichia ruminantium str Gardel |
| Clostridium botulinum B str Eklund 17B | Coronaviridae | Ehrlichia ruminantium str Welgevonden |
| Clostridium botulinum E3 str Alaska E43 | Alphacoronavirus TGEV- Purdue-P115 | Escherichia coli APEC O1 |
| Clostridium botulinum F str 230613 | Coronaviridae Bafinivirus | Escherichia coli BL21 DE3 |
| Clostridium botulinum H04402 065 | White-bream-virus | Escherichia coli B str REL606 |
| Clostridium difficile 2007855 | Coronaviridae | Escherichia coli ETEC H10407 |
| Clostridium difficile 630 | Betacoronavirus Bovine- coronavirus | Escherichia coli O104H4 str 2011C-3493 |
| Clostridium difficile B11 | Coronaviridae | Escherichia coli O157H7 str EC4115 |
| Clostridium difficile B19 | Betacoronavirus Murine- hepatitis-virus-strain-A59 | Escherichia coli O157H7 str Sakai |
| Clostridium perfringens ATCC 13124 | Coronaviridae | Escherichia coli O7K1 str CE10 |
| Clostridium perfringens SM101 | Betacoronavirus Murine- hepatitis-virus-strain-JHM | Escherichia coli O83H1 str NRG 857C |
| Clostridium perfringens str 13 | Coronaviridae | Escherichia coli str K-12 substr MG1655 |
| Clostridium symbiosum WAL- 14163 | Betacoronavirus SARS- coronavirus | Filoviridae Ebolavirus |
| Clostridium symbiosum WAL- 14673 | Coronaviridae Coronavirinae | Bundibugyo-ebolavirus |
| Clostridium tetani E88 | Munia-coronavirus-HKU13- 3514 | Filoviridae Ebolavirus Cote- d'Ivoire-ebolavirus |
| Clostridium thermocellum ATCC 27405 | Coronaviridae | Filoviridae Ebolavirus Ebola- virus-Mayinga-Zaire |
| Clostridium thermocellum DSM 1313 | Gammacoronavirus Avian- infectious-bronchitis-virus | Filoviridae Ebolavirus Reston- ebolavirus |
| Clostridium tunisiense TJ C661 | Coronaviridae | Filoviridae Ebolavirus Sudan- ebolavirus |
| Clostridium ultunense Esp | Gammacoronavirus Turkey- coronavirus | Filoviridae Marburgvirus Lake- Victoria-marburgvirus-Musoke |
| Coccidioides immitis H5384 | Coronaviridae Torovirus | Flaviviridae Alkhurma-virus |
| Coccidioides immitis RMSCC 2394 | Breda-virus | Flaviviridae Classical-swine- fever-virus |
| Coccidioides immitis RS | Coronaviridae unclassified- coronaviruses Bat- | Flaviviridae Dengue-virus 1 |
| Coccidioides posadasii C735 delta SOWgp | coronavirus-BM48-31-BGR- 2008 | Flaviviridae Dengue-virus 2 |
| Coccidioides posadasii CPA 0001 | Coronaviridae unclassified- coronaviruses Bovine- respiratory-coronavirus-AH187 | Flaviviridae Dengue-virus 3 |
| | Coronaviridae unclassified- coronaviruses Human-enteric- coronavirus-strain-4408 | Flaviviridae Dengue-virus 4 |
| | Coxiella burnetii CbuG Q212 | Flaviviridae Japanese- encephalitis-virus genome |
| | | Flaviviridae Karshi-virus |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Flaviviridae Langat-virus | Listeria monocytogenes FSL R2-561 | Poxviridae Lumpy-skin-disease-virus NI-2490 |
| Flaviviridae Louping-ill-virus | Listeria monocytogenes La111 | Poxviridae Molluscipoxvirus |
| Flaviviridae Murray-Valley-encephalitis-virus | Listeria seeligeri FSL N1-067 | Molluscum-contagiosum-virus-subtype-1 |
| Flaviviridae Omsk-hemorrhagic-fever-virus | Listeria seeligeri serovar 12b str SLCC3954 | Poxviridae Orthopoxvirus |
| Flaviviridae Powassan-virus | Listeria welshimeri serovar 6b str SLCC5334 | Camelpox-virus |
| Flaviviridae St-Louis-encephalitis-virus | Paramyxoviridae Avulavirus | Poxviridae Orthopoxvirus |
| Flaviviridae Tick-borne-encephalitis-virus | Newcastle-disease-virus-B1 | Cowpox-virus |
| Flaviviridae Usutu-virus | Paramyxoviridae Henipavirus | Poxviridae Orthopoxvirus |
| Flaviviridae West-Nile-virus | Hendra-virus | Ectromelia-virus |
| Flaviviridae Yellow-fever-virus | Paramyxoviridae Henipavirus | Poxviridae Orthopoxvirus |
| Francisella cf novicida Fx1 | Nipah-virus | Monkeypox-virus-Zaire-96-I-16 |
| Francisella noatunensis subsp orientalis str Toba 04 | Paramyxoviridae Menangle-virus | Poxviridae Orthopoxvirus |
| Francisella novicida U112 | Paramyxoviridae Morbillivirus | Taterapox-virus |
| Francisella philomiragia subsp philomiragia ATCC 25015 | Measles-virus | Poxviridae Orthopoxvirus |
| Francisella philomiragia subsp philomiragia ATCC 25017 | Paramyxoviridae Peste-des-petits-ruminants-virus | Vaccinia-virus |
| Francisella tularensis subsp holarctica F92 | Paramyxoviridae Respirivirus | Poxviridae Orthopoxvirus |
| Francisella tularensis subsp mediasiatica FSC147 | Human-parainfluenza-virus-1 | Variola-virus |
| Francisella tularensis subsp tularensis FSC198 | Paramyxoviridae Respirivirus | Poxviridae Sheeppox-virus |
| Herpesviridae Alcelaphine-herpesvirus 1 | Human-parainfluenza-virus-3 | 17077-99 |
| Herpesviridae Macacine-herpesvirus 1 | Paramyxoviridae Respirivirus | Poxviridae Suipoxvirus |
| Listeria fleischmannii LU2006-1 c88 | Sendai-virus | Swinepox-virus |
| Listeria innocua Clip11262 | Paramyxoviridae Rinderpest-virus strain-Kabete-O | Puccinia graminis f sp tritici |
| Listeria innocua FSL J1-023 | Paramyxoviridae Rubulavirus | CRL 75-36-700-3 |
| Listeria ivanovii FSL F6-596 | Human-parainfluenza-virus-2 | Ralstonia pickettii 12D |
| Listeria ivanovii subsp ivanovii PAM 55 | Paramyxoviridae Rubulavirus | Ralstonia pickettii 12J |
| Listeria marthii FSL S4-120 | Human-parainfluenza-virus-4a | Ralstonia solanacearum |
| Listeria monocytogenes 07PF0776 | Paramyxoviridae Rubulavirus | CFBP2957 |
| Listeria monocytogenes 08-5578 | Mumps-virus | Ralstonia solanacearum |
| Listeria monocytogenes 10403S | Picornaviridae Foot-and-mouth-disease-virus -type-O | CMR15 |
| Listeria monocytogenes ATCC 19117 | Picornaviridae Swine-vesicular-disease-virus strain-HK70 | Ralstonia solanacearum |
| Listeria monocytogenes Finland 1998 | Picornaviridae Swine-vesicular-disease-virus strain-NET192 | GMI1000 |
| | Poxviridae Avipoxvirus | Rathayibacter toxicus DSM 7488 |
| | Fowlpox-virus | Reoviridae African-horsesickness-virus segment-10 |
| | Poxviridae Crocodylidpoxvirus | Rhabdoviridae Vesicular-stomatitis-Indiana-virus |
| | Nile-crocodilepox-virus | Rhabdoviridae Vesicular-stomatitis-virus strain-NJ2075212NM |
| | Poxviridae Goatpox-virus | Rickettsia bellii OSU 85-389 |
| | Pellor | Rickettsia conorii str Malish 7 |
| | Poxviridae Leporipoxvirus | Rickettsia prowazekii str Breinl |
| | Myxoma-virus | Rickettsia prowazekii str RpGvF24 |
| | | Rickettsia rickettsii str Arizona |
| | | Rickettsia typhi str B9991CWPP |
| | | Rickettsiella grylli gcontig 634 |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Salmonella bongori N268-08 | Staphylococcus capitis SK14 | Xanthomonas oryzae pv |
| Salmonella bongori NCTC 12419 | Staphylococcus caprae C87 | oryzae MAFF 311018 |
| Salmonella enterica subsp arizonae serovar 62z4z23- str RSK2980 | Staphylococcus carnosus subsp carnosus TM300 | Xanthomonas oryzae pv oryzae PXO99A |
| Salmonella enterica subsp enterica serovar Dublin str CT 02021853 | Staphylococcus epidermidis ATCC 12228 | Xanthomonas vasicola pv vasculorum NCPPB 1326 scf 9767 4580 |
| Salmonella enterica subsp enterica serovar Newport str SL254 | Staphylococcus epidermidis RP62A | Yersinia aldovae ATCC 35236 |
| Salmonella enterica subsp enterica serovar Paratyphi A str ATCC 9150 | Staphylococcus equorum subsp equorum Mu2 | Yersinia bercovieri ATCC 43970 |
| Salmonella enterica subsp enterica serovar Typhimurium str LT2 | Staphylococcus haemolyticus JCSC1435 | Yersinia enterocolitica IP 10393 |
| Salmonella enterica subsp enterica serovar Typhi str Ty2 | Staphylococcus hominis SK119 | Yersinia enterocolitica IP2222 |
| Shigella boydii CDC 3083-94 | Staphylococcus hominis subsp hominis C80 | Yersinia enterocolitica subsp enterocolitica 8081 |
| Shigella boydii Sb227 | Staphylococcus lugdunensis HKU09-01 | Yersinia enterocolitica subsp palearctica 105 5R |
| Shigella dysenteriae Sd197 | Staphylococcus lugdunensis N920143 | Yersinia frederiksenii ATCC 33641 |
| Shigella flexneri 2002017 | Togaviridae Alphavirus Barmah-Forest-virus | Yersinia intermedia ATCC 29909 |
| Shigella flexneri 2a str 2457T | Togaviridae Chikungunya-virus | Yersinia kristensenii ATCC 33638 |
| Shigella flexneri 2a str 301 | Togaviridae EEEV-complex Eastern-equine-encephalitis-virus | Yersinia mollaretii ATCC 43969 |
| Shigella flexneri 5 str 8401 | Togaviridae Rubivirus Rubella-virus | Yersinia pestis A1122 |
| Shigella sonnei 53G | Togaviridae SFV-complex O-nyong-nyong-virus | Yersinia pestis Antiqua |
| Shigella sonnei Ss046 | Togaviridae Venezuelan-equine-encephalitis-virus | Yersinia pestis KIM 10 |
| Staphylococcus arlettae CVD059 SARL c230 | Togaviridae WEEV-complex Sindbis-virus | Yersinia pestis Pestoides F |
| Staphylococcus aureus 04-02981 | Togaviridae Western-equine-encephalomyelitis-virus | Yersinia pseudotuberculosis IP 31758 |
| Staphylococcus aureus 08BA02176 | Xanthomonas albilineans GPE PC73 | Yersinia pseudotuberculosis IP 32953 |
| Staphylococcus aureus subsp aureus N315 | Xanthomonas axonopodis Xac29-1 | Yersinia pseudotuberculosis PB1 |
| Staphylococcus aureus subsp aureus NCTC 8325 | Xanthomonas oryzae pv oryzae KACC 10331 | Yersinia pseudotuberculosis YPIII |
| Staphylococcus aureus subsp aureus TW20 | | Yersinia rohdei ATCC 43380 |
| Staphylococcus capitis QN1 Contig63 | | Yersinia ruckeri ATCC 29473 |

SIANN: Strain Identification by Alignment to Near Neighbors

Appendix 2: Viral Database

| | | |
|--|---|--|
| Abaca bunchy top virus DNA-C | Adoxophyes orana granulovirus | Alternanthera yellow vein virus satellite DNA beta |
| Abaca bunchy top virus DNA-M | Adoxophyes orana nucleopolyhedrovirus | Ambystoma tigrinum virus |
| Abaca bunchy top virus DNA-N | Aedes aegypti densovirus | Amsacta moorei entomopoxvirus L |
| Abaca bunchy top virus DNA-R | Aedes albopictus densovirus | Anguillid herpesvirus 1 |
| Abaca bunchy top virus DNA-S | Aedes taeniorhynchus iridescent virus | Anopheles gambiae densonucleosis virus |
| Abaca bunchy top virus segment 2 | Aeromonas phage 25 | Antheraea pernyi nucleopolyhedrovirus |
| Abalone shriveling syndrome-associated virus | Aeromonas phage 31 | Anticarsia gemmatalis nucleopolyhedrovirus |
| Abutilon Brazil virus DNA A | Aeromonas phage 44RR2.8t | Archaeal BJ1 virus |
| Abutilon Brazil virus DNA B | Aeromonas phage Aeh1 | Ateline herpesvirus 3 |
| Abutilon mosaic virus DNA A | Aeromonas phage phiO18P | Autographa californica nucleopolyhedrovirus |
| Abutilon mosaic virus DNA B | african cassava mosaic virus DNA A | Avian adeno-associated virus ATCC VR-865 |
| Acanthocystis turfacea | african cassava mosaic virus DNA B | Avian adeno-associated virus strain DA-1 |
| Chlorella virus 1 | African green monkey polyomavirus | Avian endogenous retrovirus EAV-HP |
| Acheta domesticus densovirus | African swine fever virus | Azospirillum phage Cd |
| Acholeplasma phage L2 | Ageratum enation virus | Bacillus phage 0305phi8-36 |
| Acholeplasma phage MV-L1 | Ageratum leaf Cameroon betasatellite | Bacillus phage AP50 |
| Acidianus bottle-shaped virus | Ageratum leaf curl virus-G52 | Bacillus phage B103 |
| Acidianus filamentous virus 1 | Ageratum yellow vein China virus-associated DNA beta | Bacillus phage Bam35c |
| Acidianus filamentous virus 2 | Ageratum yellow vein Chinavirus | Bacillus phage BCJA1c |
| Acidianus filamentous virus 3 | Ageratum yellow vein Hualian virus-TaiwanHsinchutom2003 DNA A | Bacillus phage Cherry |
| Acidianus filamentous virus 6 | Ageratum yellow vein Sri Lanka virus segment A | Bacillus phage Fah |
| Acidianus filamentous virus 7 | Ageratum yellow vein Taiwan virus | Bacillus phage GA-1 |
| Acidianus filamentous virus 8 | Ageratum yellow vein virusassociated DNA beta | Bacillus phage Gamma |
| Acidianus filamentous virus 9 | Ageratum yellow veinvirus | Bacillus phage GIL16c |
| Acidianus rod-shaped virus 1 | Agrotis epsilon multiple nucleopolyhedrovirus | Bacillus phage IEBH |
| Acidianus spindle-shaped virus 1 | Agrotis segetum granulovirus | Bacillus phage phi105 |
| Acidianus two-tailed virus | Agrotis segetum nucleopolyhedrovirus | Bacillus phage phi29 |
| Actinomyces phage Av-1 | Alcelaphine herpesvirus 1 | Bacillus phage SPBc2 |
| Actinoplanes phage phiAsp2 | Aleutian mink disease virus | Bacillus phage SPO1 |
| Acyrtosiphon pisum bacteriophage APSE-1 | Allamanda leaf curl virus DNA-A | Bacillus phage SPP1 |
| Adeno-associated virus-1 | Alternanthera yellow vein virus DNA-A | Bacillus phage TP21-L |
| Adeno-associated virus-2 | | Bacillus phage WBeta |
| Adeno-associated virus-3 | | Bacillus prophage phBC6A51 |
| Adeno-associated virus-4 | | Bacillus prophage phBC6A52 |
| Adeno-associated virus-5 | | Bacillus virus 1 |
| Adeno-associated virus-7 | | Bacteriophage Aaphi23 |
| Adeno-associated virus-8 | | Bacteriophage APSE-2 |
| Adoxophyes honmai NPV | | Bacteriophage PSA |
| | | Bacteriophage RB32 |
| | | Bacteroides phage B40-8 |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Banana bunchy top virus DNA C | Begomovirus-associated DNA II | Burkholderia ambifaria phage BcepF1 |
| Banana bunchy top virus DNA M | Begomovirus-associated DNA-III | Burkholderia phage Bcep 176 |
| Banana bunchy top virus DNA N | Bettongia penicillata papillomavirus 1 | Burkholderia phage Bcep1 |
| Banana bunchy top virus DNA R | Bhendi yellow vein Bhubhaneswar virus DNA-A | Burkholderia phage Bcep22 |
| Banana bunchy top virus DNA S | Bhendi yellow vein Delhi virus 2004New Delhi DNA-A | Burkholderia phage Bcep43 |
| Banana bunchy top virus DNA U3 | Bhendi yellow vein mosaic virus-associated DNA beta | Burkholderia phage Bcep781 |
| Banana streak GF virus | Bhendi yellow vein mosaic Virus | Burkholderia phage BcepB1A |
| Banana streak Mysore virus | Bitter gourd leaf curl disease-associated DNA beta | Burkholderia phage BcepC6B |
| Banana streak OL virus | BK polyomavirus | Burkholderia phage BcepGomr |
| Banana streak virus genome | Blainvillea yellow spot virus DNA-A | Burkholderia phage BcepIL02 |
| Banana streak virus strain Acuminata Vietnam | Blainvillea yellow spot virus DNA-B | Burkholderia phage BcepMu |
| Bandicoot papillomatosis carcinomatosis virus type 1 | Blattella germanica densovirus | Burkholderia phage BcepNazgul |
| Bandicoot papillomatosis carcinomatosis virus type 2 | Blueberry red ringspot virus | Burkholderia phage BcepNY3 |
| Bat adeno-associated virus YNM | Bocavirus gorillaGBoV12009 | Burkholderia phage KS10 |
| Bdellovibrio phage phiMH2K | Bombyx mandarina nucleopolyhedrovirus | Burkholderia phage KS9 |
| Beak and feather disease virus | Bombyx mori densovirus 5 | Burkholderia phage phi1026b |
| Bean calico mosaic virus DNA A | Bombyx mori NPV | Burkholderia phage phi644-2 chromosome |
| Bean calico mosaic virus DNA B | Bordetella phage BIP-1 | Burkholderia phage phiE12-2 chromosome |
| Bean dwarf mosaic virus DNA A | Bordetella phage BMP-1 | Burkholderia phage phiE125 |
| Bean dwarf mosaic virus DNA B | Bordetella phage BPP-1 | Burkholderia phage phiE202 chromosome |
| Bean golden mosaic virus DNA A | Bougainvillea spectabilis chlorotic vein-banding virus | Burkholderia phage phiE255 chromosome |
| Bean golden mosaic virus DNA B | Bovine adeno-associated virus | Burkholderia prophage phi52237 |
| Bean golden yellow mosaic virus DNA A | Bovine adenovirus A | Cabbage leaf curl virus DNA A |
| Bean golden yellow mosaic virus DNA B | Bovine adenovirus B | Cabbage leaf curl virus DNA B |
| Bean golden yellow mosaic virus DNA A | Bovine adenovirus D | Cacao swollen shoot virus |
| Bean golden yellow mosaic virus DNA B | Bovine ephemeral fever virus | California sea lion anellovirus |
| Bean yellow dwarf virus putative genes V1 | Bovine foamy virus | California sea lion polyomavirus 1 |
| Beet curly top Iran virus-K | Bovine herpesvirus 1 | Callitrichine herpesvirus 3 strain CJ0149 |
| Beet curly top virus-California Logan | Bovine herpesvirus 4 long unique region | Camelpox virus |
| Beet mild curly top virus-Worland4 | Bovine herpesvirus 5 | Campoletis sonorensis ichnovirus chromosome segment W |
| Beet severe curly top virus-Cfh | Bovine papillomavirus-1 | Campoletis sonorensis ichnovirus segment B |
| | Bovine papillomavirus 3 | Campoletis sonorensis ichnovirus segment C |
| | Bovine papular stomatitis virus | Campoletis sonorensis ichnovirus segment D |
| | Bovine parvovirus 2 | Campoletis sonorensis ichnovirus segment E |
| | Bovine Parvovirus | |
| | Bovine polyomavirus | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Campoletis sonorensis ichnovirus segment F | Capra hircus papillomavirus type 1 | Circovirus-like genome RW-E |
| Campoletis sonorensis ichnovirus segmentG2 | Cardiospermum yellow leaf curl virus satellite DNA beta | Circovirus-like genome SAR-A |
| Campoletis sonorensis ichnovirus segment G | Caretta caretta papillomavirus 1 | Circovirus-like genome SAR-B |
| Campoletis sonorensis ichnovirus segment H | Carnation etched ring virus | Citrus psorosis virus RNA1 |
| Campoletis sonorensis ichnovirus segmentI2 | Casphalia extranea densovirus | Citrus psorosis virus RNA2 |
| Campoletis sonorensis ichnovirus segment I | Cassava vein mosaic virus | Citrus psorosis virus RNA3 |
| Campoletis sonorensis ichnovirus segment J | Cauliflower mosaic virus | Citrus yellow mosaic virus |
| Campoletis sonorensis ichnovirus segment L | Cercopithecine herpesvirus 2 | Clanis bilineata nucleopolyhedrosis virus |
| Campoletis sonorensis ichnovirus segment M | Cercopithecine herpesvirus 5 strain 2715 | Clavibacter phage CMP1 |
| Campoletis sonorensis ichnovirus segment N | Cercopithecine herpesvirus 9 | Clerodendron yellow mosaic virus |
| Campoletis sonorensis ichnovirus segment O1 | Cestrum yellow leaf curling virus | Clerodendrum golden mosaic |
| Campoletis sonorensis ichnovirus segment P | Chaetoceros salsugineum DNA virus | China virus DNA A |
| Campoletis sonorensis ichnovirus segment Q | Chayote yellow mosaic virus | Clerodendrum golden mosaic |
| Campoletis sonorensis ichnovirus segment T | Chicken anemia virus | China virus DNA B |
| Campoletis sonorensis ichnovirus segment U | Chickpea chlorotic dwarf virus | Clerodendrum golden mosaic virus DNA-A |
| Campoletis sonorensis ichnovirus segment V | Chilli leaf curl disease associated sequence virion | Clerodendrum golden mosaic virus DNA-B |
| Campoletis sonorensis ichnovirus segment Z | Chilli leaf curl Multan alphasatellite | Clostridium phage 39-O |
| Campoletis superhelical segment A | Chilli leaf curl virus | Clostridium phage c-st |
| Campoletis superhelical segment aprime | Chino del tomate virus DNA A | Clostridium phage phi3626 |
| Canary circovirus | Chino del tomate virus DNA B | Clostridium phage phiC2 |
| Canarypox virus | Chlamydia phage 3 | Clostridium phage phi CD119 |
| Canine adenovirus | Chlamydia phage 4 | Clostridium phage phiCD27 |
| Canine minute virus | Chlamydia phage Chp1 | Clostridium phage phiCTP1 |
| Canine oral papillomavirus | Chlamydia phage Chp2 | Coconut foliar decay virus |
| Canine papillomavirus 2 | Chlamydia phage PhiCPG1 | Columbid circovirus |
| Canine papillomavirus 3 | Chloris striate mosaic virus | Commelina yellow mottle virus |
| Canine papillomavirus 4 | Choristoneura fumiferana DEF MNPV | Corchorus golden mosaic virus DNA-A |
| Canine papillomavirus 6 | Choristoneura fumiferana MNPV | Corchorus golden mosaic virus DNA-B |
| Canine parvovirus | Choristoneura occidentalis granulovirus | Corchorus yellow spot virus DNA A |
| | Chrysodeixis chalcites nucleopolyhedrovirus | Corchorus yellow spot virus DNA B |
| | Circovirus-like genome BBC-A | Corchorus yellow vein virus- Hoa Binh DNA A |
| | Circovirus-like genome CB-A | Corchorus yellow vein virus- Hoa Binh DNA B |
| | Circovirus-like genome CB-B | Corynebacterium phage BFK20 |
| | Circovirus-like genome RW-A | Corynebacterium phage P1201 |
| | Circovirus-like genome RW-B | Cotesia congregata bracovirus segment Circle10 |
| | Circovirus-like genome RW-C | Cotesia congregata bracovirus segment Circle11 |
| | Circovirus-like genome RW-D | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---|---|---|
| Cotesia congregata bracovirus segment Circle12 | Cotesia congregata virus segment Circle6 | Croton yellow vein virus |
| Cotesia congregata bracovirus segment Circle13 | Cotesia congregata virus segment Circle7 | Crow polyomavirus |
| Cotesia congregata bracovirus segment Circle14 | Cotesia congregata virus segment Circle8 | Cryptophlebia leucotreta granulovirus |
| Cotesia congregata bracovirus segment Circle15 | Cotton leaf crumple geminivirus DNA B | Cucurbita yellow vein virus- associated DNA beta |
| Cotesia congregata bracovirus segment Circle17 | Cotton leaf crumple virus DNA A | Cucurbit leaf crumple virus DNA A |
| Cotesia congregata bracovirus segment Circle19 | Cotton leaf curl Alabad virus | Cucurbit leaf crumple virus DNA B |
| Cotesia congregata bracovirus segment circle1 | Cotton leaf curl Bangalore virus-associated DNA beta | Culex nigripalpus NPV |
| Cotesia congregata bracovirus segment Circle20 | Cotton leaf curl Bangalore virus segment A | Culex pipiens densovirus |
| Cotesia congregata bracovirus segment Circle21 | Cotton leaf curl Burewala alphasatellite | Cyanophage PSS2 |
| Cotesia congregata bracovirus segment Circle22 | Cotton leaf curl Burewala betasatellite | Cyanophage Syn5 |
| Cotesia congregata bracovirus segment Circle23 | Cotton leaf curl Burewala virus-IndiaVehari2004 | Cycad leaf necrosis virus |
| Cotesia congregata bracovirus segment Circle25 | Cotton leaf curl Gezira alphasatellite | Cydia pomonella granulovirus |
| Cotesia congregata bracovirus segment Circle26 | Cotton leaf curl Gezira betasatellite | Cyprinid herpesvirus 3 |
| Cotesia congregata bracovirus segment circle2 | Cotton leaf curl Gezira virus | Deer papillomavirus |
| Cotesia congregata bracovirus segment Circle30 | Cotton leaf curl Kokhran virus | Deerpox virus W-1170-84 |
| Cotesia congregata bracovirus segment Circle31 | Cotton leaf curl Multan Virus | Deerpox virus W-848-83 |
| Cotesia congregata bracovirus segment Circle32 | Cotton leaf curl Multan virus satellite DNA beta | Deftia phage phiW-14 |
| Cotesia congregata bracovirus segment Circle33 | Cotton leaf curl Multan virus satellite U36-1 | Dendrolimus punctatus densovirus |
| Cotesia congregata bracovirus segment Circle35 | Cotton leaf curl Rajasthan virus segment A | Desmodium leaf distortion virus DNA A |
| Cotesia congregata bracovirus segment Circle36 | Cotton leaf curl virus- associated DNA beta | Desmodium leaf distortion virus DNA B |
| Cotesia congregata bracovirus segment circle3 | Cottontail rabbit papillomavirus | Diadromus pulchellus ascovirus 4a |
| Cotesia congregata bracovirus segment Circle4 | Cowpea severe leaf curl- associated DNA beta | Diatraea saccharalis densovirus |
| Cotesia congregata bracovirus segment Circle5 | Cowpox virus | Dicliptera yellow mottle virus DNA A |
| Cotesia congregata bracovirus segment Circle9 | Crassocephalum yellow vein virus-Jinghong | Dicliptera yellow mottle virus DNA B |
| Cotesia congregata virus segment Circle18 | Croton yellow vein mosaic alphasatellite | Digitaria streak virus |
| | Croton yellow vein mosaic virus | Dioscorea bacilliform virus |
| | Croton yellow vein mosaic Virus satellite DNA beta | Dolichos yellow mosaic virus |
| | | Dracaena mottle virus |
| | | Duck adenovirus A |
| | | Duck circovirus |
| | | Duck hepatitis B virus |
| | | East African cassava mosaic Cameroon virus DNA A |
| | | East African cassava mosaic Cameroon virus DNA B |
| | | East African cassava mosaic Kenya virus DNA A |
| | | East African cassava mosaic Kenya virus DNA B |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|---|--|
| East African cassava mosaic virus DNA A | Enterobacteria phage Min27 | Enterococcus faecalis 62 |
| East African cassava mosaic virus DNA B | Enterobacteria phage Mu | plasmid EF62pB |
| East African cassava mosaic Zanzibar virus DNA-A | Enterobacteria phage N15 | Enterococcus faecalis 62 |
| East African cassava mosaic Zanzibar virus DNA B | Enterobacteria phage N4 | plasmid EF62pC |
| Ecotropis obliqua NPV | Enterobacteria phage P1 | Enterococcus phage EF62phi |
| Ectocarpus siliculosus virus 1 | Enterobacteria phage P22 | Enterococcus phage EFAP-1 |
| Ectromelia virus | virus | Enterococcus phage phiEf11 |
| Emiliana huxleyi virus 86 | Enterobacteria phage P2 virus | Enterococcus phage phiEF24C |
| Emilia yellow vein virus-associated DNA beta | Enterobacteria phage P4 | Enterococcus phage phiFL1A |
| Emilia yellow vein virus-Fz1 | Enterobacteria phage Phi1 | Enterococcus phage phiFL2A |
| Enterobacteria phage 13a | Enterobacteria phage phiEco32 | Enterococcus phage phiFL3A |
| Enterobacteria phage 933W | Enterobacteria phage phiEcoM-GJ1 | Enterococcus phage phiFL4A |
| Enterobacteria phage alpha3 | Enterobacteria phage phiP27 | Enzootic nasal tumour virus of goats |
| Enterobacteria phage BA14 | Enterobacteria phage phiX174 sensu lato | Epiphyas postvittana NPV |
| Enterobacteria phage BP-4795 | Enterobacteria phage PRD1 | Equid herpesvirus 1 |
| Enterobacteria phage EcoDS1 | Enterobacteria phage PsP3 | Equid herpesvirus 2 |
| Enterobacteria phage EPS7 | Enterobacteria phage RB14 | Equid herpesvirus 4 |
| Enterobacteria phage epsilon15 | Enterobacteria phage RB16 | Equid herpesvirus 9 |
| Enterobacteria phage ES18 | Enterobacteria phage RB43 | Equine papillomavirus 2 |
| Enterobacteria phage Felix 01 | Enterobacteria phage RB49 | Eragrostis curvula streak virus |
| Enterobacteria phage Fels-2 | Enterobacteria phage RB51 | Eragrostis streak virus |
| Enterobacteria phage G4 | Enterobacteria phage RB69 | Erectites yellow mosaic virus DNA-A |
| Enterobacteria phage HK022 | Enterobacteria phage RTP | Erectites yellow mosaic virus satellite DNA beta |
| Enterobacteria phage HK620 | Enterobacteria phage Sf6 | Erethizon dorsatum papillomavirus type 1 |
| Enterobacteria phage HK97 | Enterobacteria phage SfV | Erinaceus europaeus papillomavirus |
| Enterobacteria phage I2-2 | Enterobacteria phage SP6 | Erwinia amylovora phage Era103 |
| Enterobacteria phage ID18 | Enterobacteria phage SSL-2009a | Erwinia phage phiEa21-4 |
| Enterobacteria phage ID2 MoscowID2001 | Enterobacteria phage ST104 | Escherichia coli bacteriophage rv5 |
| Enterobacteria phage If1 | Enterobacteria phage St-1 | Escherichia phage D108 |
| Enterobacteria phage Ike | Enterobacteria phage ST64T | Escherichia phage phiV10 |
| Enterobacteria phage IME08 | Enterobacteria phage T1 | Eupatorium vein clearing virus |
| Enterobacteria phage JK06 | Enterobacteria phage T3 | Eupatorium yellow vein associated DNA beta |
| Enterobacteria phage JS10 | Enterobacteria phage T4 | Eupatorium yellow vein virus |
| Enterobacteria phage JS98 | Enterobacteria phage T5 | Euphorbia leaf curl virus DNA A |
| Enterobacteria phage JSE | Enterobacteria phage T7 | Euphorbia yellow mosaic virus DNA A |
| Enterobacteria phage K1-5 | Enterobacteria phage TLS | Euphorbia yellow mosaic virus DNA B |
| Enterobacteria phage K1E | Enterobacteria phage VT2-Sakai | Euproctis pseudoconsersa nucleopolyhedrovirus |
| Enterobacteria phage K1F | Enterobacteria phage WA13 | European elk papillomavirus |
| Enterobacteria phage lambda | Enterobacteria phage WV8 | |
| Enterobacteria phage LKA1 | Enterobacteria phage YYZ-2008 | |
| Enterobacteria phage M13 | Enterococcus faecalis 62 chromosome | |
| | Enterococcus faecalis 62 plasmid EF62pA | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Faba bean necrotic stunt virus DNA C | Frog adenovirus 1 | Glypta fumiferanae ichnovirus segment B19 |
| Faba bean necrotic stunt virus DNA M | Frog virus 3 | Glypta fumiferanae ichnovirus segment-B1 |
| Faba bean necrotic stunt virus DNA N | Galleria mellonella densovirus | Glypta fumiferanae ichnovirus segment B20 |
| Faba bean necrotic stunt virus DNA R | Gallid herpesvirus 1 | Glypta fumiferanae ichnovirus segment B21 |
| Faba bean necrotic stunt virus DNA S | Gallid herpesvirus 2 | Glypta fumiferanae ichnovirus segment B22 |
| Faba bean necrotic stunt virus DNA U1 | Gallid herpesvirus 3 | Glypta fumiferanae ichnovirus segment B23 |
| Faba bean necrotic stunt virus DNA U2 | Gammapapillomavirus HPV127 | Glypta fumiferanae ichnovirus segment B24 |
| Faba bean necrotic stunt virus DNA U4 | Geobacillus phage GBSV1 | Glypta fumiferanae ichnovirus segment B25 |
| Faba bean necrotic yellows virus DNA 10 | Geobacillus virus E2 | Glypta fumiferanae ichnovirus segment B26 |
| Faba bean necrotic yellows virus DNA-1 | Glossina pallidipes salivary gland hypertrophy virus | Glypta fumiferanae ichnovirus segment B27 |
| Faba bean necrotic yellows virus DNA 2 | Glypta fumiferanae ichnovirus segment A10 | Glypta fumiferanae ichnovirus segment B28 |
| Faba bean necrotic yellows virus DNA 4 | Glypta fumiferanae ichnovirus segment A1 | Glypta fumiferanae ichnovirus segment B29 |
| Faba bean necrotic yellows virus DNA 5 | Glypta fumiferanae ichnovirus segment A2 | Glypta fumiferanae ichnovirus segment-B2 |
| Faba bean necrotic yellows virus DNA 7 | Glypta fumiferanae ichnovirus segment A3 | Glypta fumiferanae ichnovirus segment B30 |
| Faba bean necrotic yellows virus DNA 8 | Glypta fumiferanae ichnovirus segment A4 | Glypta fumiferanae ichnovirus segment B31 |
| Faba bean necrotic yellows virus DNA 9 | Glypta fumiferanae ichnovirus segment A5 | Glypta fumiferanae ichnovirus segment B32 |
| Faba bean necrotic yellows Virus | Glypta fumiferanae ichnovirus segment A6 | Glypta fumiferanae ichnovirus segment B33 |
| Feldmannia species virus | Glypta fumiferanae ichnovirus segment A7 | Glypta fumiferanae ichnovirus segment B34 |
| Felid herpesvirus 1 | Glypta fumiferanae ichnovirus segment A8 | Glypta fumiferanae ichnovirus segment B35 |
| Felis domesticus papillomavirus type 1 | Glypta fumiferanae ichnovirus segment A9 | Glypta fumiferanae ichnovirus segment B36 |
| Fenneropenaeus chinensis hepatopancreatic densovirus | Glypta fumiferanae ichnovirus segment B10 | Glypta fumiferanae ichnovirus segment B37 |
| Figwort mosaic virus | Glypta fumiferanae ichnovirus segment B11 | Glypta fumiferanae ichnovirus segment B38 |
| Finch circovirus | Glypta fumiferanae ichnovirus segment B12 | Glypta fumiferanae ichnovirus segment B39 |
| Finch polyomavirus | Glypta fumiferanae ichnovirus segment B13 | Glypta fumiferanae ichnovirus segment-B3 |
| Flavobacterium phage 11b | Glypta fumiferanae ichnovirus segment B14 | Glypta fumiferanae ichnovirus segment B40 |
| Fowl adenovirus A | Glypta fumiferanae ichnovirus segment B15 | |
| Fowl adenovirus D | Glypta fumiferanae ichnovirus segment B16 | |
| Fowlpox virus | Glypta fumiferanae ichnovirus segment B17 | |
| Francolinus leucoscepus papillomavirus 1 | Glypta fumiferanae ichnovirus segment B18 | |
| Fringilla coelebs papillomavirus | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---|---|---|
| Glypta fumiferanae ichnovirus segment B41 | Glypta fumiferanae ichnovirus segment B64 | Glypta fumiferanae ichnovirus segment C7 |
| Glypta fumiferanae ichnovirus segment B42 | Glypta fumiferanae ichnovirus segment B65 | Glypta fumiferanae ichnovirus segment C8 |
| Glypta fumiferanae ichnovirus segment B43 | Glypta fumiferanae ichnovirus segment-B6 | Glypta fumiferanae ichnovirus segment C9 |
| Glypta fumiferanae ichnovirus segment B44 | Glypta fumiferanae ichnovirus segment B7 | Glypta fumiferanae ichnovirus segment D1 |
| Glypta fumiferanae ichnovirus segment B45 | Glypta fumiferanae ichnovirus segment B8 | Glypta fumiferanae ichnovirus segment D2 |
| Glypta fumiferanae ichnovirus segment B46 | Glypta fumiferanae ichnovirus segment B9 | Glypta fumiferanae ichnovirus segment D3 |
| Glypta fumiferanae ichnovirus segment B47 | Glypta fumiferanae ichnovirus segment C10 | Glypta fumiferanae ichnovirus segment D4 |
| Glypta fumiferanae ichnovirus segment B48 | Glypta fumiferanae ichnovirus segment C11 | Glypta fumiferanae ichnovirus segment D5 |
| Glypta fumiferanae ichnovirus segment B49 | Glypta fumiferanae ichnovirus segment C12 | Glypta fumiferanae ichnovirus segment D6 |
| Glypta fumiferanae ichnovirus segment-B4 | Glypta fumiferanae ichnovirus segment C13 | Glypta fumiferanae ichnovirus segment D7 |
| Glypta fumiferanae ichnovirus segment B50 | Glypta fumiferanae ichnovirus segment C14 | Glypta fumiferanae ichnovirus segment E1 |
| Glypta fumiferanae ichnovirus segment B51 | Glypta fumiferanae ichnovirus segment C15 | Goatpox virus Pellor |
| Glypta fumiferanae ichnovirus segment B52 | Glypta fumiferanae ichnovirus segment C16 | Goose circovirus |
| Glypta fumiferanae ichnovirus segment B53 | Glypta fumiferanae ichnovirus segment C17 | Goose hemorrhagic polyomavirus |
| Glypta fumiferanae ichnovirus segment B54 | Glypta fumiferanae ichnovirus segment C18 | Goose parvovirus |
| Glypta fumiferanae ichnovirus segment B55 | Glypta fumiferanae ichnovirus segment C19 | Gossypium darwinii symptomless alphasatellite DNA-alpha |
| Glypta fumiferanae ichnovirus segment B56 | Glypta fumiferanae ichnovirus segment-C1 | Gossypium darwinii symptomless virus DNA-A |
| Glypta fumiferanae ichnovirus segment B57 | Glypta fumiferanae ichnovirus segment C20 | Gossypium davidsonii symptomless alphasatellite DNA-alpha-B |
| Glypta fumiferanae ichnovirus segment B58 | Glypta fumiferanae ichnovirus segment C21 | Gossypium mustelinum symptomless alphasatellite DNA-alpha-B |
| Glypta fumiferanae ichnovirus segment B59 | Glypta fumiferanae ichnovirus segment C22 | Gossypium punctatum mild leaf curl virus DNA A |
| Glypta fumiferanae ichnovirus segment-B5 | Glypta fumiferanae ichnovirus segment-C2 | Gossypium punctatum mild leaf curl virus DNA B |
| Glypta fumiferanae ichnovirus segment B60 | Glypta fumiferanae ichnovirus segment C3 | Ground squirrel hepatitis virus |
| Glypta fumiferanae ichnovirus segment B61 | Glypta fumiferanae ichnovirus segment C4 | Gryllus bimaculatus nudivirus |
| Glypta fumiferanae ichnovirus segment B62 | Glypta fumiferanae ichnovirus segment C5 | Gull circovirus |
| Glypta fumiferanae ichnovirus segment B63 | Glypta fumiferanae ichnovirus segment C6 | Haemophilus phage HP1 |
| | | Haemophilus phage HP2 |
| | | Haloarcula hispanica pleomorphic virus 1 |
| | | Haloarcula phage SH1 |
| | | Halomonas phage phiHAP-1 |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--------------------------------|------------------------------|--------------------------------|
| Halorubrum phage HF2 | Human adenovirus F | Hyperthermophilic Archaeal |
| Halorubrum pleomorphic virus | Human bocavirus 1 | Virus 1 |
| 1 | Human bocavirus 2 | Hyperthermophilic Archaeal |
| Halovirus HF1 | Human bocavirus 3 | Virus 2 |
| Hamster polyomavirus | Human bocavirus 4 | Hyphantria cunea |
| Helicoverpa armigera | Human erythrovirus V9 | nucleopolyhedrovirus |
| granulovirus | Human herpesvirus 1 | Hyposoter fugitivus ichnovirus |
| Helicoverpa armigera | Human herpesvirus 2 | segment A1 |
| multiple nucleopolyhedrovirus | Human herpesvirus 3 | Hyposoter fugitivus ichnovirus |
| Helicoverpa armigera-NPV | Human herpesvirus 4 | segment A2 |
| Helicoverpa armigera NPV | Human herpesvirus 5 strain | Hyposoter fugitivus ichnovirus |
| NNg1 | Merlin | segment A3 |
| Helicoverpa armigera | Human herpesvirus 7 | Hyposoter fugitivus ichnovirus |
| nucleopolyhedrovirus G4 | Human herpesvirus 8 | segment B10 |
| Helicoverpa zea SNPV | Human papillomavirus-18 | Hyposoter fugitivus ichnovirus |
| Heliopsis virescens ascovirus | Human papillomavirus 1 | segment B11 |
| 3e | Human papillomavirus-2 | Hyposoter fugitivus ichnovirus |
| Hepatitis B virus | Human papillomavirus 54 | segment B12 |
| Heron hepatitis B virus | Human papillomavirus-5 | Hyposoter fugitivus ichnovirus |
| His1 virus | Human papillomavirus type | segment B13 |
| His2 virus | 101 | Hyposoter fugitivus ichnovirus |
| Hollyhock leaf crumple virus | Human papillomavirus type | segment B14 |
| Honeysuckle yellow vein | 103 | Hyposoter fugitivus ichnovirus |
| beta-JapanFukui2001 | Human papillomavirus type | segment B15 |
| Honeysuckle yellow vein | 108 | Hyposoter fugitivus ichnovirus |
| mosaic beta- | Human papillomavirus type- | segment B16 |
| JapanMiyazaki2001 | 10 | Hyposoter fugitivus ichnovirus |
| Honeysuckle yellow vein | Human papillomavirus type 16 | segment B17 |
| mosaic disease associated | Human papillomavirus type 26 | Hyposoter fugitivus ichnovirus |
| satellite DNA beta-Ibaraki | Human papillomavirus type 32 | segment B18 |
| Honeysuckle yellow vein | Human papillomavirus type 34 | Hyposoter fugitivus ichnovirus |
| mosaic Virus | Human papillomavirus type 41 | segment-B1 |
| Honeysuckle yellow vein | Human papillomavirus type 48 | Hyposoter fugitivus ichnovirus |
| mosaic virus-Kagoshima | Human papillomavirus type 49 | segment B2 |
| Honeysuckle yellow vein | Human papillomavirus type-4 | Hyposoter fugitivus ichnovirus |
| mosaic virus satellite DNA | Human papillomavirus type 50 | segment B3 |
| beta | Human papillomavirus type 53 | Hyposoter fugitivus ichnovirus |
| Honeysuckle yellow vein virus- | Human papillomavirus type 60 | segment B4 |
| UK1 | Human papillomavirus type 63 | Hyposoter fugitivus ichnovirus |
| Horsegram yellow mosaic | Human papillomavirus type | segment B5 |
| Virus DNA B | 6b | Hyposoter fugitivus ichnovirus |
| Horsegram yellow mosaic | Human papillomavirus type 7 | segment B6 |
| virus | Human papillomavirus type 88 | Hyposoter fugitivus ichnovirus |
| Horseradish curly top virus | Human papillomavirus type 90 | segment B7 |
| Human adenovirus 54 | Human papillomavirus type 92 | Hyposoter fugitivus ichnovirus |
| Human adenovirus AHuman | Human papillomavirus type 96 | segment B8 |
| adenovirus B1 | Human papillomavirus type-9 | Hyposoter fugitivus ichnovirus |
| Human adenovirus B2 | Human parvovirus B19 | segment B9 |
| Human adenovirus C | Human T-lymphotropic virus 1 | Hyposoter fugitivus ichnovirus |
| Human adenovirus D | Human T-lymphotropic virus 4 | segment C10 |
| Human adenovirus E | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Hyposoter fugitivus ichnovirus segment C11 | Hyposoter fugitivus ichnovirus segment D4 | Lactobacillus phage Lc-Nu |
| Hyposoter fugitivus ichnovirus segment C12 | Hyposoter fugitivus ichnovirus segment D5 | Lactobacillus phage LL-H |
| Hyposoter fugitivus ichnovirus segment C13 | Hyposoter fugitivus ichnovirus segment D6 | Lactobacillus phage LP65 |
| Hyposoter fugitivus ichnovirus segment C14 | Hyposoter fugitivus ichnovirus segment D7 | Lactobacillus phage Lrm1 |
| Hyposoter fugitivus ichnovirus segment C15 | Hyposoter fugitivus ichnovirus segment D8 | Lactobacillus phage Lv-1 |
| Hyposoter fugitivus ichnovirus segment C16 | Hyposoter fugitivus ichnovirus segment D9 | Lactobacillus phage phiAT3 |
| Hyposoter fugitivus ichnovirus segment C17 | Hyposoter fugitivus ichnovirus segment E1 | Lactobacillus phage phig1e |
| Hyposoter fugitivus ichnovirus segment C18 | Hyposoter fugitivus ichnovirus segment E2 | Lactobacillus phage phiJL-1 |
| Hyposoter fugitivus ichnovirus segment C19 | Hyposoter fugitivus ichnovirus segment G1 | Lactobacillus prophage Lj928 |
| Hyposoter fugitivus ichnovirus segment-C1 | Ictalurid herpesvirus 1 strain Auburn 1 | Lactobacillus prophage Lj965 |
| Hyposoter fugitivus ichnovirus segment C20 | Indian cassava mosaic virus DNA A | Lactobacillus prophage phiadh |
| Hyposoter fugitivus ichnovirus segment-C2 | Indian cassava mosaic virus DNA B | Lactococcus phage 1706 |
| Hyposoter fugitivus ichnovirus segment C3 | Infectious hypodermal and hematopoietic necrosis virus | Lactococcus phage 4268 |
| Hyposoter fugitivus ichnovirus segment C4 | Infectious spleen and kidney necrosis virus | Lactococcus phage 712 |
| Hyposoter fugitivus ichnovirus segment C5 | Invertebrate iridescent virus 6 | Lactococcus phage |
| Hyposoter fugitivus ichnovirus segment C6 | Iodobacteriophage phiPLPE | ascphi28 |
| Hyposoter fugitivus ichnovirus segment C7 | Ipomoea yellow vein virus | Lactococcus phage blBB29 |
| Hyposoter fugitivus ichnovirus segment C8 | Jatropha leaf curl virus DNA A | Lactococcus phage blL170 |
| Hyposoter fugitivus ichnovirus segment C9 | Jatropha yellow mosaic India virus DNA-A | Lactococcus phage blL67 |
| Hyposoter fugitivus ichnovirus segment D10 | JC polyomavirus | Lactococcus phage BK5-T |
| Hyposoter fugitivus ichnovirus segment D11 | Junonia coenia densovirus | Lactococcus phage c2 |
| Hyposoter fugitivus ichnovirus segment D12 | Kalanchoe top-spotting virus | Lactococcus phage jj50 |
| Hyposoter fugitivus ichnovirus segment-D1 | Kenaf leaf curl virus DNA A | Lactococcus phage KSY1 |
| Hyposoter fugitivus ichnovirus segment D2 | KI polyomavirus Stockholm 60 | Lactococcus phage P008 |
| Hyposoter fugitivus ichnovirus segment D3 | Klebsiella phage KP15 | Lactococcus phage P087 |
| | Klebsiella phage KP32 | Lactococcus phage phiLC3 |
| | Klebsiella phage KP34 | Lactococcus phage Q54 |
| | Kluyvera phage Kvp1 | Lactococcus phage r1t |
| | Kudzu mosaic virus DNA-A | Lactococcus phage sk1 |
| | Kudzu mosaic virus DNA-B | Lactococcus phage TP901-1 |
| | Lactobacillus johnsonii prophage Lj771 | Lactococcus phage Tuc2009 |
| | Lactobacillus phage A2 | Lactococcus phage ul36 |
| | Lactobacillus phage KC5a | Lactococcus prophage |
| | Lactobacillus phage Lb338-1 | blL285 |
| | | Lactococcus prophage |
| | | blL286 |
| | | Lactococcus prophage |
| | | blL309 |
| | | Lactococcus prophage |
| | | blL310 |
| | | Lactococcus prophage |
| | | blL311 |
| | | Lactococcus prophage |
| | | blL312 |
| | | Lamium leaf distortion associated virus |
| | | Leucania separata nuclear polyhedrosis virus |
| | | Leucas zeylanica yellow vein virus satellite DNA beta |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|---|---|
| Lindernia anagallis yellow vein virus DNA-A | JamaicaWissadulaAugust Town DNA B | Mastomys natalensis papillomavirus |
| Lindernia anagallis yellow vein virus satellite DNA beta | Macroptilium mosaic Puerto Rico virus DNA A | Melanoplus sanguinipes entomopoxvirus |
| Listeria phage A006 | Macroptilium mosaic Puerto Rico virus DNA B | Meleagrid herpesvirus 1 |
| Listeria phage A118 | Macroptilium yellow mosaic Florida virus DNA A | Melon chlorotic leaf curl virus DNA A |
| Listeria phage A500 | Macroptilium yellow mosaic Florida virus DNA B | Melon chlorotic mosaic virus-associated alphasatellite |
| Listeria phage A511 | Macroptilium yellow mosaic virus DNA A | Melon chlorotic mosaic virus DNA-A |
| Listeria phage B025 | Macroptilium yellow mosaic virus DNA B | Melon chlorotic mosaic virus DNA-B |
| Listeria phage B054 | Maize streak virus-ASouth Africa | Merkel cell polyomavirus |
| Listeria phage P35 | Malachra yellow vein mosaic virus-associated satellite DNA beta | Merremia mosaic virus DNA A |
| Listeria phage P40 | Mal de Rio Cuarto virus segment 9 | Merremia mosaic virus DNA B |
| Listonella phage phiHSIC | Malvastrum leaf curl Guangdong virus | Mesta yellow vein mosaic Bahaich virus-IndiaBahaich2007 DNA A |
| Loofa yellow mosaic virus DNA A | Malvastrum leaf curl virus-associated defective DNA beta | Mesta yellow vein mosaic virus-associated DNA beta |
| Loofa yellow mosaic virus DNA B | Malvastrum leaf curl virus-G87 | Mesta yellow vein mosaic virus DNA-A |
| Lucky bamboo bacilliform virus | Malvastrum yellow mosaic virus-associated DNA 1 | Methanobacterium phage psiM2 |
| Ludwigia leaf distortion betasatellite | Malvastrum yellow mosaic virus DNA-A | Methanothermobacter prophage psiM100 |
| IndiaAmadalavalasaHibiscus2007 | Malvastrum yellow mosaic virus satellite DNA beta | Microbacterium phage Min1 |
| Ludwigia yellow vein virus-associated DNA beta | Malvastrum yellow vein Baoshan virus DNA-A | Microcystis phage Ma-LMM01 |
| Ludwigia yellow vein virus DNA-A | Malvastrum yellow vein-virus | Microplitis demolitor bracovirus segment A |
| Luffa begomovirus associated DNA beta | Malvastrum yellow vein virus satellite DNA beta | Microplitis demolitor bracovirus segment B |
| Luffa puckering and leaf distortion-associated DNA beta | Malvastrum yellow vein Yunnan-virus | Microplitis demolitor bracovirus segment C |
| Lull virus | Malvastrum yellow vein Yunnan virus satellite DNA beta | Microplitis demolitor bracovirus segment D |
| Lumpy skin disease virus NI-2490 | Mamestra configurata NPV-A | Microplitis demolitor bracovirus segment E |
| Lymantria dispar MNPV | Mamestra configurata NPV-B | Microplitis demolitor bracovirus segment F |
| Lymantria xylinia MNPV | Mannheimia phage phiMHa1 | Microplitis demolitor bracovirus segment G |
| Lymphocystis disease virus 1 | Maruca vitrata MNPV | Microplitis demolitor bracovirus segment H |
| Lymphocystis disease virus-isolate China | Mastomys coucha papillomavirus 2 | Microplitis demolitor bracovirus segment I |
| Macacine herpesvirus 1 | | Microplitis demolitor bracovirus segment J |
| Macacine herpesvirus 3 | | Microplitis demolitor bracovirus segment K |
| Macacine herpesvirus 4 | | |
| Macacine herpesvirus 5 | | |
| Macaque simian foamy virus | | |
| Macroptilium golden mosaic virus- | | |
| JamaicaWissadulaAugust Town DNA A | | |
| Macroptilium golden mosaic virus- | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---|---|------------------------------------|
| Microplitis demolitor bracovirus segment L | Mungbean yellow mosaic India virus DNA A | Mycobacterium phage Che9c |
| Microplitis demolitor bracovirus segment M | Mungbean yellow mosaic India virus DNA B | Mycobacterium phage Che9d |
| Microplitis demolitor bracovirus segment N | Mungbean yellow mosaic virus DNA A | Mycobacterium phage Cjw1 |
| Microplitis demolitor bracovirus segment O | Mungbean yellow mosaic virus DNA B | Mycobacterium phage Cooper |
| Milk vetch dwarf virus segment 10 | Murid herpesvirus 1 | Mycobacterium phage Corndog |
| Milk vetch dwarf virus segment 11 | Murid herpesvirus 2 | Mycobacterium phage CrimD |
| Milk vetch dwarf virus segment-1 | Murid herpesvirus 4 | Mycobacterium phage D29 |
| Milk vetch dwarf virus segment 2 | Murine adenovirus 3 | Mycobacterium phage DD5 |
| Milk vetch dwarf virus segment 3 | Murine adenovirus A | Mycobacterium phage ET08 |
| Milk vetch dwarf virus segment 4 | Murine pneumotropic virus | Mycobacterium phage Fruitloop |
| Milk vetch dwarf virus segment 5 | Murine polyomavirus | Mycobacterium phage Giles |
| Milk vetch dwarf virus segment 6 | Murine type C retrovirus | Mycobacterium phage Gumball |
| Milk vetch dwarf virus segment 7 | Musca domestica salivary gland hypertrophy virus | Mycobacterium phage Halo |
| Milk vetch dwarf virus segment 8 | Muscovy duck circovirus | Mycobacterium phage Jasper |
| Milk vetch dwarf virus segment 9 | Muscovy duck parvovirus | Mycobacterium phage KBG |
| Mimosa yellow leaf curl virus- associated DNA 1 | Mus musculus papillomavirus type 1 | Mycobacterium phage Konstantine |
| Mimosa yellow leaf curl virus DNA-A | Mycobacterium phage 244 | Mycobacterium phage Kostya |
| Mimosa yellow leaf curl virus satellite DNA beta | Mycobacterium phage Adjutor | Mycobacterium phage L5 |
| Minute virus of mice | Mycobacterium phage Angel | Mycobacterium phage LeBron |
| Mirabilis mosaic virus | Mycobacterium phage angelica | Mycobacterium phage Llij |
| Miscanthus streak virus-91 | Mycobacterium phage Ardmore | Mycobacterium phage Lockley |
| Molluscum contagiosum virus subtype 1 | Mycobacterium phage Barnyard | Mycobacterium phage Myrna |
| Monkeypox virus Zaire-96-I-16 | Mycobacterium phage Bethlehem | Mycobacterium phage Nigel |
| Morganella phage MmP1 | Mycobacterium phage Boomer | Mycobacterium phage Omega |
| Mouse mammary tumor virus | Mycobacterium phage BPs | Mycobacterium phage Orion |
| Mouse parvovirus 1 | Mycobacterium phage Brujita | Mycobacterium phage Pacc40 |
| Mouse parvovirus 2 | Mycobacterium phage Butterscotch | Mycobacterium phage PBI1 |
| Mouse parvovirus 3 | Mycobacterium phage Bxb1 | Mycobacterium phage Peaches |
| Mouse parvovirus 4 | Mycobacterium phage Bxz1 | Mycobacterium phage PG1 |
| Mouse parvovirus 5 | Mycobacterium phage Bxz2 | Mycobacterium phage Phaedrus |
| Mulard duck circovirus | Mycobacterium phage Cali | Mycobacterium phage Phlyer |
| | Mycobacterium phage Catera | Mycobacterium phage Pipefish |
| | Mycobacterium phage Chah | Mycobacterium phage PLOT |
| | Mycobacterium phage Che12 | Mycobacterium phage PMC |
| | Mycobacterium phage Che8 | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|-------------------------------|-------------------------------|----------------------------------|
| Mycobacterium phage Porky | Okra yellow mosaic Mexico | Pelargonium vein banding virus |
| Mycobacterium phage | virus DNA B | |
| Predator | Okra yellow vein disease | Penaeus merguensis |
| Mycobacterium phage | associated sequence virion | densovirus |
| Pukovnik | Okra yellow vein mosaic virus | Penaeus monodon |
| Mycobacterium phage | Old World harvest mouse | hepatopancreatic parvovirus |
| Gyrzula | papillomavirus | Pepper curly top virus |
| Mycobacterium phage | Orangutan polyomavirus | Pepper golden mosaic virus |
| Ramsey | Orf virus | DNA A |
| Mycobacterium phage Rizal | Orgyia leucostigma NPV | Pepper golden mosaic virus |
| Mycobacterium phage | Orgyia pseudotsugata MNPV | DNA B |
| Rosebush | Oryctes rhinoceros virus | Pepper huasteco yellow vein |
| Mycobacterium phage | Ostreid herpesvirus 1 | virus DNA A |
| ScottMcG | Ostreococcus tauri virus 1 | Pepper huasteco yellow vein |
| Mycobacterium phage Solon | Ostreococcus virus OsV5 | virus DNA B |
| Mycobacterium phage Spud | Ovine adenovirus A | Pepper leaf curl Bangladesh |
| Mycobacterium phage TM4 | Ovine adenovirus D | virus segment A component |
| Mycobacterium phage Troll4 | Ovine herpesvirus 2 | Pepper leaf curl virus DNA-A |
| Mycobacterium phage | Ovine papillomavirus-1 | Pepper leaf curl virus satellite |
| Tweety | Panicum streak virus-Karino | DNA beta |
| Mycobacterium phage U2 | Panine herpesvirus 2 strain | Pepper leaf curl Yunnan virus |
| Mycobacterium phage | Heberling | satellite DNA beta |
| Wildcat | Papaya leaf curl China virus- | Pepper leaf curl Yunnan virus- |
| Mycoplasma phage MAV1 | G8 | YN323 |
| Mycoplasma phage P1 | Papaya leaf curl China virus | Pepper yellow dwarf virus- |
| Mycoplasma phage phiMFV1 | satellite DNA beta | New Mexico |
| Myotis polyomavirus VM-2008 | Papaya leaf curl Guandong | Pepper yellow leaf curl |
| Mythimna loreyi densovirus | virus-GD2 DNA A | Indonesia virus DNA-A |
| Myxococcus phage Mx8 | Papaya leaf curl virus- | Pepper yellow leaf curl |
| Myxoma virus | associated DNA beta | Indonesia virus DNA-B |
| Myzus persicae densovirus | Papaya leaf curl-virus | Pepper yellow vein Mali virus |
| Nanovirus-like particle | Papiine herpesvirus 2 | Periplaneta fuliginosa |
| Natralba phage PhiCh1 | Paramecium bursaria | densovirus |
| Neodiprion abietis NPV | Chlorella virus 1 | Petunia vein clearing virus |
| Neodiprion lecontei NPV | Paramecium bursaria | Phage cdtI |
| Neodiprion sertifer NPV | Chlorella virus AR158 | Phage Gfsy-1 |
| Oat dwarf virus | Paramecium bursaria | Phage Gfsy-2 |
| OkLCV satDNA 10 | Chlorella virus FR483 | Phage phiJL001 |
| Okra leaf curl disease | Paramecium bursaria | Phocoena spinipinnis |
| associated DNA 1 | Chlorella virus NY2A | papillomavirus |
| Okra leaf curl Mali virus | Parvovirus H1 | Phormidium phage Pf-WMP3 |
| satellite DNA beta | Passionfruit severe leaf | Phormidium phage Pf-WMP4 |
| Okra leaf curl virus-Cameroon | distortion virus DNA-A | Phthorimaea operculella |
| Okra mottle virus-Brazilokra | Passionfruit severe leaf | granulovirus |
| DNA A | distortion virus DNA-B | Pieris rapae granulovirus |
| Okra mottle virus-Brazilokra | Pasteurella phage F108 | Planaria asexual strain- |
| DNA B | Peanut chlorotic streak virus | specific virus-like element |
| Okra yellow crinkle virus | Pedilanthus leaf curl virus- | type 1 large DNA segment |
| segment A | Pedilanthus | Planaria asexual strain- |
| Okra yellow mosaic Mexico | PakistanMultan2004 | specific virus-like element |
| virus DNA A | | type 1 small DNA segment |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---|---|---|
| Planococcus citri densovirus | Pseudomonas phage LBL3 | Ranid herpesvirus 1 strain |
| Plutella xylostella granulovirus | Pseudomonas phage LIT1 | McKinnell |
| Plutella xylostella multiple nucleopolyhedrovirus | Pseudomonas phage LKD16 | Ranid herpesvirus 2 strain |
| Polyomavirus HPyV6 | Pseudomonas phage LMA2 | ATCC VR-568 |
| Polyomavirus HPyV7 | Pseudomonas phage LUZ19 | Rauscher murine leukemia virus |
| Porcine adenovirus C | Pseudomonas phage LUZ24 | Raven circovirus |
| Porcine circovirus 1 | Pseudomonas phage LUZ7 | RD114 retrovirus |
| Porcine circovirus 2 | Pseudomonas phage M6 | Reticuloendotheliosis virus |
| Porcine endogenous retrovirus E | Pseudomonas phage MP22 | Rhesus monkey papillomavirus |
| Porcine parvovirus | Pseudomonas phage MP29 | Rhizobium phage 16-3 |
| Potato apical leaf curl disease-associated satellite DNA beta | Pseudomonas phage MP38 | Rhodothermus phage RM378 |
| Potato yellow mosaic | Pseudomonas phage PA11 | Rhynchosia golden mosaic virus DNA A |
| Panama virus DNA A | Pseudomonas phage PAJU2 | Rhynchosia golden mosaic virus DNA B |
| Potato yellow mosaic | Pseudomonas phage PaP2 | Rhynchosia golden mosaic Yucatan virus DNA A |
| Panama virus DNA B | Pseudomonas phage PaP3 | Rhynchosia golden mosaic Yucatan virus DNA B |
| Potato yellow mosaic Trinidad virus DNA A | Pseudomonas phage PB1 | Rice tungro bacilliform virus |
| Potato yellow mosaic Trinidad virus DNA B | Pseudomonas phage Pf1 | Roseobacter phage SIO1 |
| Potato yellow mosaic virus DNA A | Pseudomonas phage Pf3 | Roseophage DSS3P2 |
| Potato yellow mosaic virus DNA B | Pseudomonas phage phi-2 | Roseophage EE36P1 |
| Prochlorococcus phage P-SSM4 | Pseudomonas phage phiCTX | Rosss goose hepatitis B virus |
| Propionibacterium phage B5 | Pseudomonas phage phikF77 | Rousettus aegyptiacus papillomavirus type 1 |
| Propionibacterium phage PA6 | Pseudomonas phage phikMV | Rudbeckia flower distortion virus |
| Pseudaletia unipuncta granulovirus | Pseudomonas phage phiKZ | Saccharum streak virus |
| Pseudoalteromonas phage PM2 | Pseudomonas phage PT2 | Saimiriine herpesvirus 2 |
| Pseudocowpox virus | Pseudomonas phage PT5 | Salmonella enterica bacteriophage SE1 |
| Pseudomonas phage 119X | Pseudomonas phage SN | Salmonella phage c341 |
| Pseudomonas phage 14-1 | Pseudomonas phage YuA | Salmonella phage E1 |
| Pseudomonas phage 201phi2-1 | Psittacid herpesvirus 1 | Salmonella phage epsilon34 |
| Pseudomonas phage 73 | Psittacus erithacus timneh papillomavirus | Salmonella phage Fels-1 |
| Pseudomonas phage B3 | Pumpkin yellow mosaic | Salmonella phage phiSG-JL2 |
| Pseudomonas phage D3112 | Malaysia virus DNA A | Salmonella phage SETP3 |
| Pseudomonas phage-D3 | Pyrobaculum spherical virus | Salmonella phage ST64B |
| Pseudomonas phage DMS3 | Pyrococcus abyssi virus 1 | Sclerotinia sclerotiorum hypovirulence associated DNA virus 1 |
| Pseudomonas phage EL | Rabbit fibroma virus | Sea turtle tornovirus 1 |
| Pseudomonas phage F10 | Rabbit oral papillomavirus | Senecio yellow mosaic virus |
| Pseudomonas phage F116 | Rachiplusia ou MNPV | Sheeppox virus 17077-99 |
| Pseudomonas phage F8 | Radish leaf curl virus satellite DNA beta | Sheldgoose hepatitis B virus |
| Pseudomonas phage gh-1 | Radish leaf curl virus segment A | Shigella phage phiSboM-AG3 |
| | Ralstonia phage p12J | Shrimp white spot syndrome virus |
| | Ralstonia phage phiRSA1 | |
| | Ralstonia phage RSB1 | |
| | Ralstonia phage RSL1 | |
| | Ralstonia phage RSM1 | |
| | Ralstonia phage RSM3 | |
| | Ralstonia phage RSS1 | |
| | Ramie mosaic virus DNA-A | |
| | Ramie mosaic virus DNA-B | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|---|---|
| Sida golden mosaic Costa Rica virus DNA A | Sida yellow vein Vietnam virus satellite DNA beta | Spiroplasma phage SVTS2 |
| Sida golden mosaic Costa Rica virus DNA B | Sida yellow vein virus DNA A | Spodoptera exigua MNPV |
| Sida golden mosaic Florida virus-Malvastrum DNA-A | Sida yellow vein virus DNA B | Spodoptera frugiperda ascovirus 1a |
| Sida golden mosaic Florida virus-Malvastrum DNA-B | Sida yellow vein virus satellite DNA beta | Spodoptera frugiperda MNPV virus |
| Sida golden mosaic Honduras virus DNA A | Siegesbeckia yellow vein Guangxi virus | Spodoptera litura granulovirus |
| Sida golden mosaic Honduras virus DNA B | Siegesbeckia yellow vein virus-GD13-associated DNA beta | Spodoptera litura NPV |
| Sida golden mosaic virus DNA-A | Siegesbeckia yellow vein virus GD13 | Spodoptera litura nucleopolyhedrovirus II |
| Sida golden mosaic virus DNA-B | Simian adenovirus 1 | Sputnik virophage |
| Sida golden mottle virus DNA-A | Simian adenovirus 3 | Squash leaf curl China virus-B DNA-A |
| Sida golden mottle virus DNA-B | Simian foamy virus 3 | Squash leaf curl China virus-B DNA B |
| Sida golden mottle virus DNA-B | Simian immunodeficiency virus SIV-mnd 2 | Squash leaf curl Philippines virus segment A |
| Sida leaf curl virus-associated DNA 1 | Simian retrovirus 4 | Squash leaf curl Philippines virus segment B |
| Sida leaf curl virus-associated DNA beta | Simian T-cell lymphotropic virus 6 | Squash leaf curl virus A component DNA |
| Sida leaf curl-virus | Simian T-lymphotropic virus 1 | Squash leaf curl virus B component DNA |
| Sida leaf curl virus satellite DNA beta | Simian T-lymphotropic virus 3 | Squash leaf curl Yunnan virus |
| Sida micrantha mosaic virus segment A | Simian virus 40 | Squash mild leaf curl virus-Imperial Valley DNA A |
| Sida micrantha mosaic virus segment B | Singapore grouper iridovirus | Squash mild leaf curl virus-Imperial Valley DNA B |
| Sida mosaic Sinaloa virus DNA A | Sinorhizobium phage PBC5 | Squash yellow mild mottle virus DNA B |
| Sida mosaic Sinaloa virus DNA B | Small anellovirus 1 | Squirrel monkey polyomavirus |
| Sida mottle virus | Small anellovirus 2 | Sri Lankan cassava mosaic virus DNA A |
| Sida yellow mosaic virus-China-associated DNA beta | Snake parvovirus 1 | Sri Lankan cassava mosaic virus DNA B |
| Sida yellow mosaic-virus | Snow goose hepatitis B virus | Stachytarpheta leaf curl virus |
| Sida yellow mosaic Yucatan virus DNA A | Sodalis phage phiSG1 | Staphylococcus phage 11 |
| Sida yellow mosaic Yucatan virus DNA B | Sodalis phage SO-1 | Staphylococcus phage 187 |
| Sida yellow vein disease associated DNA 1 | South African cassava mosaic virus DNA A | Staphylococcus phage 2638A |
| Sida yellow vein Madurai virus | South African cassava mosaic virus DNA B | Staphylococcus phage 29 |
| Sida yellow vein Vietnam virus-associated DNA 1 | Soybean chlorotic blotch virus DNA A | Staphylococcus phage 37 |
| Sida yellow vein Vietnam virus DNA-A | Soybean chlorotic blotch virus DNA B | Staphylococcus phage 3A |
| | Soybean chlorotic mottle virus | Staphylococcus phage 42e |
| | Soybean crinkle leaf virus | Staphylococcus phage 44AHJD |
| | Spilanthes yellow vein virus DNA-A | Staphylococcus phage 47 |
| | Spinach curly top virus | Staphylococcus phage 52A |
| | Spiroplasma kunkelii virus SkV1 CR2-3x | Staphylococcus phage 53 |
| | Spiroplasma phage 1-C74 | Staphylococcus phage 55 |
| | Spiroplasma phage 1-R8A2B | Staphylococcus phage 66 |
| | Spiroplasma phage 4 | Staphylococcus phage 69 |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | | | |
|-------------------------------------|-------------------------------|----------|--|----------|
| Staphylococcus phage 71 | Staphylococcus | prophage | Streptococcus | prophage |
| Staphylococcus phage 77 | phi 12 | | MM1 | |
| Staphylococcus phage 80alpha | Staphylococcus | prophage | Streptococcus | pyogenes |
| Staphylococcus phage 85 | phi 13 | | phage 315.3 | |
| Staphylococcus phage 88 | Staphylococcus | prophage | Streptomyces phage mu16 | |
| Staphylococcus phage 92 | phiN315 | | Streptomyces phage phiBT1 | |
| Staphylococcus phage 96 | Staphylococcus | prophage | Streptomyces phage phiC31 | |
| Staphylococcus phage CNPH82 | phiPV83 | | Streptomyces phage | |
| Staphylococcus phage EW | Staphylococcus | prophage | phiSASD1 | |
| Staphylococcus phage G1 | PVL | | Streptomyces phage VWB | |
| Staphylococcus phage K | Staphylococcus | prophage | Stx1 converting phage | |
| Staphylococcus phage P954 | tp310-1 | | Stx2-converting phage 1717 | |
| Staphylococcus phage PH15 | Staphylococcus | prophage | Stx2-converting phage 86 | |
| Staphylococcus phage phi2958PVL | tp310-2 | | Stx2 converting phage-I | |
| Staphylococcus phage phiETA2 | Staphylococcus | prophage | Stx2 converting phage II | |
| Staphylococcus phage phiETA3 | tp310-3 | | Subterranean clover stunt | |
| Staphylococcus phage-phiETA | Starling circovirus | | virus DNA 1 | |
| Staphylococcus phage phiMR11 | Stenotrophomonas | phage | Subterranean clover stunt | |
| Staphylococcus phage phiMR25 | phiSMA9 | | virus DNA 2 | |
| Staphylococcus phage phiNM1 | Stenotrophomonas phage S1 | | Subterranean clover stunt | |
| Staphylococcus phage phiNM3 | Strawberry vein banding virus | | virus DNA 3 | |
| Staphylococcus phage phiP68 | Streptococcus phage 2972 | | Subterranean clover stunt | |
| Staphylococcus phage phiPVL108 | Streptococcus phage 5093 | | virus DNA 4 | |
| Staphylococcus phage phiPVL-CN125 | Streptococcus phage 7201 | | Subterranean clover stunt | |
| Staphylococcus phage phiSauS-IPLA35 | Streptococcus phage 858 | | virus DNA 5 | |
| Staphylococcus phage phiSauS-IPLA88 | Streptococcus phage Abc2 | | Subterranean clover stunt | |
| Staphylococcus phage phiSLT | Streptococcus phage | | virus DNA 6 | |
| Staphylococcus phage PT1028 | ALQ13.2 | | Subterranean clover stunt | |
| Staphylococcus phage ROSA | Streptococcus phage C1 | | virus DNA 7 | |
| Staphylococcus phage SAP-26 | Streptococcus phage Cp-1 | | Subterranean clover stunt | |
| Staphylococcus phage SAP2 | Streptococcus phage DT1 | | virus DNA 8 | |
| Staphylococcus phage Twort | Streptococcus phage M102 | | Sugarcane bacilliform IM virus | |
| Staphylococcus phage X2 | Streptococcus phage O1205 | | Sugarcane bacilliform Mor virus | |
| | Streptococcus phage P9 | | Sugarcane bacilliform virus | |
| | Streptococcus phage PH10 | | Sugarcane streak Egypt virus-Giza | |
| | Streptococcus phage PH15 | | Sugarcane streak Reunion virus | |
| | Streptococcus phage phi3396 | | Sugarcane streak virus-Natal | |
| | Streptococcus phage Sfi11 | | Suid herpesvirus 1 | |
| | Streptococcus phage Sfi19 | | Sulfolobus islandicus | |
| | Streptococcus phage Sfi21 | | filamentous virus | |
| | Streptococcus phage SM1 | | Sulfolobus islandicus rod-shaped virus 1 | |
| | Streptococcus phage SMP | | Sulfolobus islandicus rod-shaped virus 2 | |
| | Streptococcus prophage 315.1 | | Sulfolobus spindle-shaped virus 4 | |
| | Streptococcus prophage 315.2 | | Sulfolobus spindle-shaped virus 5 | |
| | Streptococcus prophage 315.5 | | | |
| | Streptococcus prophage 315.6 | | | |
| | Streptococcus prophage EJ-1 | | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | | |
|--|----------------|--|--|
| Sulfolobus virus 6 | spindle-shaped | Tobacco curly shoot-virus | Tomato leaf curl Cebu virus DNA-A |
| Sulfolobus virus 7 | spindle-shaped | Tobacco leaf curl disease associated sequence virion | Tomato leaf curl China virus-G32 |
| Sulfolobus icosahedral virus 2 | turreted | Tobacco leaf curl Japan virus | Tomato leaf curl Cotabato virus DNA-A |
| Sulfolobus icosahedral-virus | turreted | Tobacco leaf curl Thailand virus | Tomato leaf curl Ghana virus segment A |
| Sulfolobus virus 1 | | Tobacco leaf curl virus-associated DNA beta | Tomato leaf curl Guangdong virus DNA-A |
| Sulfolobus virus 2 | | Tobacco leaf curl Yunnan virus associated DNA 1 | Tomato leaf curl Guangxi virus |
| Sulfolobus virus Kamchatka 1 | | Tobacco leaf curl Yunnan virus satellite DNA beta | Tomato leaf curl Gujarat virus-Varanasi segment A |
| Sulfolobus virus Ragged Hills | | Tobacco leaf curl Yunnan virus-Y136 | Tomato leaf curl Gujarat virus-Varanasi segment B |
| Sulfolobus virus STSV1 | | Tobacco leaf curl Zimbabwe virus | Tomato leaf curl Hainan virus |
| Sunn hemp leaf distortion virus DNA-A | | Tobacco vein clearing virus | Tomato leaf curl Hsinchu virus-TaiwanHsinchu2005 DNA A |
| Sus scrofa papillomavirus type 1 | | Tobacco yellow dwarf virus | Tomato leaf curl Iran virus |
| Sweetpotato badnavirus B | | Tomato begomovirus satellite DNA beta | Tomato leaf curl Java virus-Ageratum satellite DNA |
| Sweet potato leaf curl Bengal virus-IndiaWest Bengal2008 segment A | | Tomato chino La Paz virus segment A | Tomato leaf curl Java-virus |
| Sweet potato leaf curl Canary virus | | Tomato chlorotic mottle virus DNA A | Tomato leaf curl Joydebpur beta virus |
| Sweet potato leaf curl Georgia virus | | Tomato chlorotic mottle virus DNA B | Tomato leaf curl Joydebpur virus DNA-A |
| Sweet potato leaf curl Lanzarote virus | | Tomato common mosaic virus DNA-A | Tomato leaf curl Karnataka virus-associated DNA beta DNA-A |
| Sweet potato leaf curl Spain virus | | Tomato common mosaic virus DNA-B | Tomato leaf curl Karnataka-virus |
| Sweet potato leaf curl virus | | Tomato curly stunt virus | Tomato leaf curl Kerala virus |
| Swinepox virus | | Tomato golden mosaic virus DNA A | Tomato leaf curl Kumasi virus segment A |
| Synechococcus phage P60 | | Tomato golden mosaic virus DNA B | Tomato leaf curl Laos virus |
| Synechococcus phage S-PM2 | | Tomato golden mottle virus DNA A | Tomato leaf curl Malaysia virus |
| Synechococcus phage S-RSM4 | | Tomato golden mottle virus DNA B | Tomato leaf curl Mali virus |
| Synechococcus phage syn9 | | Tomato golden mottle virus DNA B | Tomato leaf curl Mayotte virus |
| Tanapox virus | | Tomato leaf curl Arusha virus DNA-A | Tomato leaf curl Mindanao virus DNA-A |
| Taro bacilliform virus | | Tomato leaf curl Bangalore virus-Ban5 satellite DNA beta | Tomato leaf curl New Delhi virus-associated DNA beta |
| Taterapox virus | | Tomato leaf curl Bangalore-virus | Tomato leaf curl New Delhi virus DNA A |
| Thalassomonas phage BA3 | | Tomato leaf curl Bangladesh virus | Tomato leaf curl New Delhi virus DNA B |
| Thermoproteus tenax spherical virus 1 | | Tomato leaf curl Cameroon virus- | Tomato leaf curl Nigeria virus-Nigeria2006 |
| Thermus phage IN93 | | CameroonBueaOkra2008 | Tomato leaf curl Pakistan virus associated DNA 1 |
| Thermus phage P23-45 | | | |
| Thermus phage P23-77 | | | |
| Thermus phage P74-26 | | | |
| Thermus phage phiYS40 | | | |
| Tobacco curly shoot virus associated DNA 1 | | | |
| Tobacco curly shoot virus-associated DNA beta | | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---|---|---|
| Tomato leaf curl Pakistan virus segment A | Tomato mottle Taino virus DNA B | Tomato yellow spot virus DNA-B |
| Tomato leaf curl Palampur virus | Tomato mottle virus DNA A | Tomato yellow vein streak virus DNA-A |
| Tomato leaf curl Patna virus DNA-A | Tomato mottle virus DNA B | Tomato yellow vein streak virus DNA-B |
| Tomato leaf curl Philippines virus | Tomato pseudo-curly top virus DNA A | Torque teno canis virus |
| Tomato leaf curl Philippine virus satellite DNA beta | Tomato rugose mosaic virus DNA B | Torque teno douroucouli virus |
| Tomato leaf curl Pune virus | Tomato severe leaf curl virus DNA A | Torque teno felis virus |
| Tomato leaf curl Seychelles virus | Tomato severe rugose virus DNA A | Torque teno midi virus 1 |
| Tomato leaf curl Sinaloa virus DNA A | Tomato severe rugose virus DNA B | Torque teno midi virus 2 |
| Tomato leaf curl Sinaloa virus DNA B | Tomato yellow dwarf disease associated satellite DNA beta-Kochi virus | Torque teno mini virus 1 |
| Tomato leaf curl Sri Lanka virus | Tomato yellow leaf curl China virus associated DNA 1 | Torque teno mini virus 2 |
| Tomato leaf curl Sudan virus-Gezira | Tomato yellow leaf curl China-virus | Torque teno mini virus 3 |
| Tomato leaf curl Sulawesi virus DNA-A | Tomato yellow leaf curl Guangdong virus DNA-A | Torque teno mini virus 4 |
| Tomato leaf curl Taiwan virus | Tomato yellow leaf curl Indonesia virus-Lembang | Torque teno mini virus 5 |
| Tomato leaf curl Togo virus-Togo2006 | Tomato yellow leaf curl Kanchanaburi virus DNA A | Torque teno mini virus 6 |
| Tomato leaf curl Vietnam virus DNA A | Tomato yellow leaf curl Kanchanaburi virus DNA B | Torque teno mini virus 7 |
| Tomato leaf curl virus-associated DNA beta | Tomato yellow leaf curl Mali virus-associated DNA beta | Torque teno mini virus 8 |
| Tomato leaf curl-virus | Tomato yellow leaf curl Thailand virus associated DNA 1 | Torque teno mini virus 9 |
| Tomato leaf curl virus-Pune-associated DNA beta DNA-A | Tomato yellow leaf curl Thailand virus DNA A | Torque teno sus virus 1 |
| Tomato mild mosaic virus DNA-A | Tomato yellow leaf curl Thailand virus DNA B | Torque teno tamarin virus |
| Tomato mild mosaic virus DNA-B | Tomato yellow leaf curl Vietnam virus DNA-A | Torque teno virus 10 |
| Tomato mild yellow leaf curl Aragua virus DNA A | Tomato yellow leaf curl Vietnam virus satellite DNA beta | Torque teno virus 12 |
| Tomato mild yellow leaf curl Aragua virus DNA B | Tomato yellow leaf curl virus-associated DNA beta | Torque teno virus 14 |
| Tomato mosaic Havana virus DNA A | Tomato yellow margin leaf curl virus DNA A | Torque teno virus 15 |
| Tomato mosaic Havana virus DNA B | Tomato yellow margin leaf curl virus DNA B | Torque teno virus 16 |
| Tomato mosaic leaf curl virus DNA A | Tomato yellow spot virus DNA-A | Torque teno virus 19 |
| Tomato mosaic leaf curl virus DNA B | | Torque teno virus-1 |
| Tomato mottle Taino virus DNA A | | Torque teno virus 25 |
| | | Torque teno virus 26 |
| | | Torque teno virus 27 |
| | | Torque teno virus 28 |
| | | Torque teno virus 3 |
| | | Torque teno virus 4 |
| | | Torque teno virus 6 |
| | | Torque teno virus 7 |
| | | Torque teno virus 8 |
| | | Trichodysplasia spinulosa-associated polyomavirus |
| | | Trichoplusia ni ascovirus 2c |
| | | Trichoplusia ni SNPV |
| | | Tupaiaid herpesvirus 1 |
| | | Turkey adenovirus A |
| | | Turnip curly top virus |
| | | TYLCCNV-Y322 satellite DNA beta |
| | | Urochloa streak virus |
| | | Vaccinia virus |
| | | Variola virus |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---------------------------------------|---|---|
| Velvet bean severe mosaic virus DNA A | Vibrio phage VP2 | Xanthomonas phage phiL7 |
| Velvet bean severe mosaic virus DNA B | Vibriophage VP4 | Xanthomonas phage Xop411 |
| Vernonia yellow vein betasatellite | Vibrio phage VP5 | Xanthomonas phage Xp10 |
| Vernonia yellow vein virus DNA-A | Vibrio phage VP882 | Xanthomonas phage Xp15 |
| Vibrio phage fs1 | Vibrio phage VP93 | Xenopus laevis endogenous retrovirus Xen1 |
| Vibrio phage fs2 | Vibrio phage VpV262 | Xestia c-nigrum granulovirus |
| Vibrio phage K139 | Vibrio phage VSK | Xylella phage Xfas53 |
| Vibrio phage kappa | Watermelon chlorotic stunt virus DNA A | Yaba-like disease virus |
| Vibrio phage KSF-1phi | Watermelon chlorotic stunt virus DNA B | Yaba monkey tumor virus |
| Vibrio phage KVP40 | Wheat dwarf virus | Yersinia pestis phage phiA1122 |
| Vibrio phage N4 | Wissadula golden mosaic St Thomas Virus DNA A | Yersinia phage Berlin |
| Vibrio phage VEJphi | Wissadula golden mosaic St Thomas Virus DNA B | Yersinia phage L-413C |
| Vibrio phage Vf12 | Woodchuck hepatitis virus | Yersinia phage phiYeO3-12 |
| Vibrio phage VfO3K6 | WU Polyomavirus | Yersinia phage PY54 |
| Vibrio phage VfO4K68 | Xanthomonas phage Cf1c | Zinnia leaf curl disease associated sequence virion |
| Vibrio phage VGJphi | Xanthomonas phage OP1 | Zinnia leaf curl virus-associated DNA beta |
| Vibrio phage VHML | Xanthomonas phage OP2 | |

SIANN: Strain Identification by Alignment to Near Neighbors

Appendix 3: Bacterial Database

| | | |
|--|--|---|
| Acaryochloris marina MBIC11017 | Actinosynnema mirum DSM 43827 | Anaerolinea thermophila UNI 1 |
| Acetobacter pasteurianus IFO 3283 01 | Aerococcus urinae ACS 120 V Col10a | Anaeromyxobacter dehalogenans 2CP 1 |
| Acetohalobium arabaticum DSM 5501 | Aeromonas hydrophila subsp hydrophila ATCC 7966 | Anaeromyxobacter sp Fw109 5 |
| Acholeplasma laidlawii PG 8A | Aeromonas salmonicida subsp salmonicida A449 | Anaplasma centrale str Israel |
| Achromobacter xylosoxidans A8 | Aeromonas veronii B565 | Anaplasma marginale str Florida |
| Acidaminococcus fermentans DSM 20731 | Aeropyrum pernix K1 | Anaplasma phagocytophilum HZ |
| Acidaminococcus intestini RyC MR95 | Aggregatibacter actinomycetemcomitans D11S 1 | Anoxybacillus flavithermus WK1 |
| Acidianus hospitalis W1 | Aggregatibacter aphrophilus NJ8700 | Aquifex aeolicus VF5 |
| Acidilobus saccharovorans 345 15 | Agrobacterium fabrum str C58 | Arcanobacterium haemolyticum DSM 20595 |
| Acidimicrobium ferrooxidans DSM 10331 | Agrobacterium radiobacter K84 | Archaeoglobus fulgidus DSM 4304 |
| Acidiphilium cryptum JF 5 | Agrobacterium sp H13 3 | Archaeoglobus profundus DSM 5631 |
| Acidiphilium multivorum AIU301 | Agrobacterium vitis S4 | Archaeoglobus veneficus SNP6 |
| Acidithiobacillus caldus SM 1 | Akkermansia muciniphila ATCC BAA 835 | Arcobacter butzleri RM4018 |
| Acidithiobacillus ferrivorans SS3 | Alcanivorax borkumensis SK2 | Arcobacter nitrofigilis DSM 7299 |
| Acidithiobacillus ferrooxidans ATCC 23270 | Alicyclophilus denitrificans BC | Arcobacter sp L |
| Acidobacterium capsulatum ATCC 51196 | Alicyclobacillus acidocaldarius subsp acidocaldarius DSM 446 | Aromatoleum aromaticum EbN1 |
| Acidothermus cellulolyticus 11B | Aliivibrio salmonicida LFI1238 | Arthrobacter arilaitensis Re117 |
| Acidovorax avenae subsp avenae ATCC 19860 | Alkalilimnicola ehrlichii MLHE 1 | Arthrobacter aurescens TC1 |
| Acidovorax citrulli AAC00 1 | Alkaliphilus metalliredigens QYMF | Arthrobacter chlorophenolicus A6 |
| Acidovorax ebreus TPSY | Alkaliphilus oremlandii OhLAs | Arthrobacter phenanthrenivorans Sphe3 |
| Acidovorax sp JS42 | Allochrochromatium vinosum DSM 180 | Arthrobacter sp FB24 |
| Aciduliprofundum boonei T469 | Alteromonas macleodii str Deep ecotype | Aster yellows witches broom phytoplasma AYWB |
| Acinetobacter baumannii 1656 2 | Alteromonas sp SN2 | Asticcacaulis excentricus CB 48 |
| Acinetobacter calcoaceticus PHEA 2 | Aminobacterium colombiense DSM 12261 | Atopobium parvulum DSM 20469 |
| Acinetobacter oleivorans DR1 | Ammonifex degensii KC4 | Azoarcus sp BH72 |
| Acinetobacter sp ADP1 | Amycolatopsis mediterranei U32 | Azorhizobium caulinodans ORS 571 |
| Actinobacillus pleuropneumoniae serovar 5b str L20 | Amycolicococcus subflavus DQS3 9A1 | Azospirillum sp B510 |
| Actinobacillus succinogenes 130Z | Anabaena variabilis ATCC 29413 | Azotobacter vinelandii DJ |
| Actinoplanes sp SE50 110 | Anaerococcus prevotii DSM 20548 | Bacillus amyloliquefaciens DSM 7 |
| | | Bacillus anthracis str Ames |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Bacillus atrophaeus 1942 | Bifidobacterium animalis | Burkholderia gladioli BSR3 |
| Bacillus cellulosilyticus DSM 2522 | subsp animalis ATCC 25527 | Burkholderia glumae BGR1 |
| Bacillus cereus 03BB102 | Bifidobacterium bifidum S17 | Burkholderia mallei ATCC 23344 |
| Bacillus clausii KSM K16 | Bifidobacterium breve ACS 071 V Sch8b | Burkholderia multivorans ATCC 17616 |
| Bacillus coagulans 2 6 | Bifidobacterium dentium Bd1 | Burkholderia phymatum STM815 |
| Bacillus cytotoxicus NVH 391 98 | Bifidobacterium longum DJO10A | Burkholderia phytofirmans PsJN |
| Bacillus halodurans C 125 | Blattabacterium sp Blattella germanica str Bge | Burkholderia pseudomallei 668 |
| Bacillus licheniformis DSM 13 ATCC 14580 | Bordetella avium 197N | Burkholderia rhizoxinica HKI 454 |
| Bacillus megaterium DSM 319 | Bordetella bronchiseptica RB50 | Burkholderia sp 383 |
| Bacillus mycoides DSM 2048 | Bordetella parapertussis 12822 | Burkholderia thailandensis E264 |
| Bacillus pseudofirmus OF4 | Bordetella pertussis CS | Burkholderia vietnamiensis G4 |
| Bacillus pumilus SAFR 032 | Bordetella petrii DSM 12804 | Burkholderia xenovorans LB400 |
| Bacillus selenitireducens MLS10 | Borrelia afzelii PKo | Butyrivibrio proteoclasticus B316 |
| Bacillus subtilis BSn5 | Borrelia bissetii DN127 | Caldicellulosiruptor bescii DSM 6725 |
| Bacillus thuringiensis serovar berliner ATCC 10792 | Borrelia burgdorferi B31 | Caldicellulosiruptor hydrothermalis 108 |
| Bacillus weihenstephanensis KBAB4 | Borrelia duttonii Ly | Caldicellulosiruptor kristjanssonii 177R1B |
| Bacteriovorax marinus SJ | Borrelia garinii PBI | Caldicellulosiruptor kronotskyensis 2002 |
| Bacteroides fragilis YCH46 | Borrelia hermsii DAH | Caldicellulosiruptor lactoaceticus 6A |
| Bacteroides helcogenes P 36 108 | Borrelia recurrentis A1 | Caldicellulosiruptor obsidiansis OB47 |
| Bacteroides salanitronis DSM 18170 | Borrelia turicatae 91E135 | Caldicellulosiruptor owensensis OL |
| Bacteroides thetaiotaomicron VPI 5482 | Brachybacterium faecium DSM 4810 | Caldicellulosiruptor saccharolyticus DSM 8903 |
| Bacteroides vulgatus ATCC 8482 | Brachyspira hyodysenteriae WA1 | Calditerrivibrio nitroreducens DSM 19672 |
| Bartonella bacilliformis KC583 | Brachyspira intermedia PWS A | Caldivirga maquilingensis IC 167 |
| Bartonella clarridgeiae 73 | Brachyspira murdochii DSM 12563 | Campylobacter concisus 13826 |
| Bartonella grahamii as4aup | Brachyspira pilosicoli 95 1000 | Campylobacter curvus 52592 |
| Bartonella henselae str Houston 1 | Bradyrhizobium japonicum USDA 110 | Campylobacter fetus subsp fetus 82 40 |
| Bartonella quintana str Toulouse | Bradyrhizobium sp BTAi1 | Campylobacter hominis ATCC BAA 381 |
| Bartonella tribocorum CIP 105476 | Brevibacillus brevis NBRC 100599 | Campylobacter jejuni subsp jejuni 81116 |
| Baumannia cicadellinicola str Hc Homalodisca coagulata | Brevundimonas subvibrioides ATCC 15264 | |
| Bdellovibrio bacteriovorus HD100 | Brucella abortus A13334 | |
| Beijerinckia indica subsp indica ATCC 9039 | Brucella canis ATCC 23365 | |
| Beutenbergia cavernae DSM 12333 | Brucella melitensis bv 1 str 16M | |
| Bifidobacterium adolescentis ATCC 15703 | Brucella microti CCM 4915 | |
| | Brucella ovis ATCC 25840 | |
| | Brucella pinnipedialis B2 94 | |
| | Brucella suis 1330 | |
| | Buchnera aphidicola str APS | |
| | Acyrtosiphon pisum | |
| | Burkholderia ambifaria AMMD | |
| | Burkholderia cenocepacia HI2424 | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|---|---|
| Campylobacter lari RM2100 | Candidatus Solibacter usitatus Ellin6076 | Chlorobium phaeobacteroides DSM 266 |
| Candidatus Accumulibacter phosphatis clade IIA str UW 1 | Candidatus Sulcia muelleri GWSS | Chlorobium phaeovibrioides DSM 265 |
| Candidatus Amoebophilus asiaticus 5a2 | Candidatus Tremblaya princeps PCIT | Chlorobium tepidum TLS |
| Candidatus Arthromitus sp SFB mouse Japan | Candidatus Vesicomysocius okutanii HA | Chloroflexus aggregans DSM 9485 |
| Candidatus Azobacteroides pseudotrichonymphae genomovar CFP2 | Candidatus Zinderia insecticola CARI | Chloroflexus aurantiacus J 10 fl |
| Candidatus Blochmannia floridanus | Capnocytophaga canimorsus Cc5 | Chloroflexus sp Y 400 fl |
| Candidatus Carsonella ruddii PV | Capnocytophaga ochracea DSM 7271 | Chloroherpeton thalassium ATCC 35110 |
| Candidatus Chloracidobacterium thermophilum B | Carboxydotherrus hydrogenoformans Z 2901 | Chromobacterium violaceum ATCC 12472 |
| Candidatus Desulforudis audaxviator MP104C | Carnobacterium sp 17 4 | Chromohalobacter salexigens DSM 3043 |
| Candidatus Hamiltonella defensa 5AT Acyrthosiphon pisum | Catenulispora acidiphila DSM 44928 | Citrobacter koseri ATCC BAA 895 |
| Candidatus Hodgkinia cicadicola Dsem | Caulobacter crescentus CB15 | Citrobacter rodentium ICC168 |
| Candidatus Korarchaeum cryptofilum OPF8 | Caulobacter segnis ATCC 21756 | Clavibacter michiganensis subsp michiganensis NCPPB 382 |
| Candidatus Koribacter versatilis Ellin345 | Caulobacter sp K31 | Clostridium acetobutylicum ATCC 824 |
| Candidatus Liberibacter asiaticus str psy62 | Cellulomonas fimi ATCC 484 | Clostridium beijerinckii NCIMB 8052 |
| Candidatus Methyloirabilis oxyfera | Cellulomonas flavigena DSM 20109 | Clostridium botulinum A str ATCC 3502 |
| Candidatus Midichloria mitochondrii IricVA | Cellulophaga algicola DSM 14237 | Clostridium cellulolyticum H10 |
| Candidatus Moranella endobia PCIT | Cellulophaga lytica DSM 7489 | Clostridium cellulovorans 743B |
| Candidatus Nitrospira defluvii | Cellvibrio gilvus ATCC 13127 | Clostridium difficile 630 |
| Candidatus Pelagibacter sp IMCC9063 | Cellvibrio japonicus Ueda107 | Clostridium kluyveri DSM 555 |
| Candidatus Phytoplasma australiense | Cenarchaeum symbiosum A | Clostridium lentocellum DSM 5427 |
| Candidatus Protochlamydia amoebophila UWE25 | Chelativorans sp BNC1 | Clostridium ljungdahlii DSM 13528 |
| Candidatus Puniceispirillum marinum IMCC1322 | Chitinophaga pinensis DSM 2588 | Clostridium novyi NT |
| Candidatus Riesia pediculicola USDA | Chlamydia muridarum Nigg | Clostridium perfringens ATCC 13124 |
| Candidatus Ruthia magnifica str Cm Calyptogena magnifica | Chlamydia trachomatis 434 Bu | Clostridium phytofermentans ISDg |
| | Chlamydophila abortus S26 3 | Clostridium saccharolyticum WM1 |
| | Chlamydophila caviae GPIC | Clostridium sp SY8519 |
| | Chlamydophila felis Fe C 56 | Clostridium sticklandii DSM 519 |
| | Chlamydophila pecorum E58 | Clostridium tetani E88 |
| | Chlamydophila pneumoniae CWL029 | Clostridium thermocellum ATCC 27405 |
| | Chlamydophila psittaci 6BC | Collimonas fungivorans Ter331 |
| | Chlorobaculum parvum NCIB 8327 | Colwellia psychrerythraea 34H |
| | Chlorobium chlorochromatii CaD3 | |
| | Chlorobium limicola DSM 245 | |
| | Chlorobium luteolum DSM 273 | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Comamonas testosteroni CNB 2 | Dechloromonas aromatica RCB | Desulfovibrio aespoeensis Aspo 2 |
| Conexibacter woesei DSM 14684 | Deferibacter desulfuricans SSM1 | Desulfovibrio africanus str Walvis Bay |
| Coprothermobacter proteolyticus DSM 5265 | Dehalococcoides ethenogenes 195 | Desulfovibrio alaskensis G20 |
| Coralimargarita akajimensis DSM 45221 | Dehalococcoides sp BAV1 | Desulfovibrio desulfuricans ND132 |
| Coriobacterium glomerans PW2 | Dehalogenimonas lykanthroporepellens BL DC 9 | Desulfovibrio magneticus RS 1 |
| Corynebacterium aurimucosum ATCC 700975 | Deinococcus deserti VCD115 | Desulfovibrio salexigens DSM 2638 |
| Corynebacterium diphtheriae NCTC 13129 | Deinococcus geothermalis DSM 11300 | Desulfovibrio vulgaris RCH1 |
| Corynebacterium efficiens YS 314 | Deinococcus maricopensis DSM 21211 | Desulfurispirillum indicum S5 |
| Corynebacterium glutamicum ATCC 13032 | Deinococcus proteolyticus MRP | Desulfurivibrio alkaliphilus AHT2 |
| Corynebacterium jeikeium K411 | Deinococcus radiodurans R1 | Desulfurobacterium thermolithotrophum DSM 11699 |
| Corynebacterium kroppenstedtii DSM 44385 | Delftia acidovorans SPH 1 | Desulfurococcus kamchatkensis 1221n |
| Corynebacterium pseudotuberculosis FRC41 | Delftia sp Cs1 4 | Desulfurococcus mucosus DSM 2162 |
| Corynebacterium resistens DSM 45100 | Denitrovibrio acetiphilus DSM 12809 | Dichelobacter nodosus VCS1703A |
| Corynebacterium ulcerans 809 | Desulfarculus baarsii DSM 2075 | Dickeya dadantii 3937 |
| Corynebacterium urealyticum DSM 7109 | Desulfatibacillum alkenivorans AK 01 | Dickeya zeae Ech1591 |
| Corynebacterium variabile DSM 44702 | Desulfatibacillum hafniense Y51 | Dictyoglomus thermophilum H 6 12 |
| Coxiella burnetii RSA 493 | Desulfobacca acetoxidans DSM 11109 | Dictyoglomus turgidum DSM 6724 |
| Croceibacter atlanticus HTCC2559 | Desulfobacterium autotrophicum HRM2 | Dinoroseobacter shibae DFL 12 |
| Cronobacter sakazakii ATCC BAA 894 | Desulfobulbus propionicus DSM 2032 | Dyadobacter fermentans DSM 18053 |
| Cronobacter turicensis z3032 | Desulfococcus oleovorans Hxd3 | Edwardsiella ictaluri 93 146 |
| Cryptobacterium curtum DSM 15641 | Desulfohalobium retbaense DSM 5692 | Edwardsiella tarda EIB202 |
| Cupriavidus metallidurans CH34 | Desulfomicrobium baculatum DSM 4028 | Eggerthella lenta DSM 2243 |
| Cupriavidus necator N 1 | Desulfotalea psychrophila LSv54 | Eggerthella sp YY7918 |
| Cupriavidus taiwanensis LMG 19424 | Desulfotomaculum acetoxidans DSM 771 | Ehrlichia canis str Jake |
| cyanobacterium UCYN A | Desulfotomaculum carboxydovorans CO 1 SRB | Ehrlichia chaffeensis str Arkansas |
| Cyanothece sp ATCC 51142 | Desulfotomaculum kuznetsovii DSM 6115 | Ehrlichia ruminantium str Welgevonden |
| Cyclobacterium marinum DSM 745 | Desulfotomaculum reducens MI 1 | Elusimicrobium minutum Pei191 |
| Cytophaga hutchinsonii ATCC 33406 | Desulfotomaculum ruminis DSM 2154 | Enterobacter aerogenes KCTC 2190 |
| | | Enterobacter asburiae LF7a |
| | | Enterobacter cloacae subsp cloacae ATCC 13047 |
| | | Enterobacter sp 638 |
| | | Enterococcus faecalis V583 |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---------------------------------------|
| Erwinia amylovora ATCC 49946 | Francisella tularensis subsp holarctica LVS | Granulicella mallensis MP5ACTX8 |
| Erwinia billingiae Eb661 | Frankia alni ACN14a | Granulicella tundricola |
| Erwinia pyrifoliae DSM 12163 | Frankia sp Ccl3 | Haemophilus ducreyi 35000HP |
| Erwinia sp Ejp617 | Frankia symbiont of Datisca glomerata | Haemophilus influenzae 10810 |
| Erwinia tasmaniensis Et1 99 | Fusobacterium nucleatum subsp nucleatum ATCC 25586 | Haemophilus parainfluenzae T3T1 |
| Erysipelothrix rhusiopathiae | Gallibacterium anatis UMN179 | Haemophilus parasuis SH0165 |
| Erythrobacter litoralis HTCC2594 | Gallionella capsiferiformans ES 2 | Haemophilus somnus 129PT |
| Escherichia coli O157:H7 str Sakai | gamma proteobacterium HdN1 | Hahella chejuensis KCTC 2396 |
| Escherichia fergusonii ATCC 35469 | Gardnerella vaginalis 409 05 | Halalkalicoccus jeotgali B3 |
| Ethanoligenens harbinense YUAN 3 | Gemmatimonas aurantiaca T 27 | Halanaerobium hydrogeniformans |
| Eubacterium eligens ATCC 27750 | Geobacillus kaustophilus HTA426 | Haliangium ochraceum DSM 14365 |
| Eubacterium limosum KIST612 | Geobacillus sp C56 T3 | Haliscomenobacter hydrossis DSM 1100 |
| Eubacterium rectale ATCC 33656 | Geobacillus thermodenitrificans NG80 2 | Haloarcula hispanica ATCC 33960 |
| Exiguobacterium sibiricum 255 15 | Geobacillus thermoglucosidasius C56 YS93 | Haloarcula marismortui ATCC 43049 |
| Exiguobacterium sp AT1b | Geobacillus thermoleovorans CCB US3 UF5 | Halobacterium sp NRC 1 |
| Ferrimonas balearica DSM 9799 | Geobacter bemidjiensis Bem | Haloferax volcanii DS2 |
| Ferroglobus placidus DSM 10642 | Geobacter daltonii FRC 32 | Halogeometricum borinquense DSM 11551 |
| Fervidobacterium nodosum Rt17 B1 | Geobacter lovleyi SZ | Halomicrobium mukohataei DSM 12286 |
| Fibrobacter succinogenes subsp succinogenes S85 | Geobacter metallireducens GS 15 | Halomonas elongata DSM 2581 |
| Filifactor alocis ATCC 35896 | Geobacter sp M18 | halophilic archaeon DL31 |
| Finegoldia magna ATCC 29328 | Geobacter sulfurreducens PCA | Halopiger xanaduensis SH 6 |
| Flavobacteriaceae bacterium 3519 10 | Geobacter uraniireducens Rf4 | Haloquadratum walsbyi C23 |
| Flavobacterium branchiophilum FL 15 | Geodermatophilus obscurus DSM 43160 | Halorhabdus utahensis DSM 12940 |
| Flavobacterium columnare ATCC 49512 | Glaciecola nitratireducens FR1064 | Halorhodospira halophila SL1 |
| Flavobacterium johnsoniae UW101 | Glaciecola sp 4H 3 7YE 5 | Halorubrum lacusprofundi ATCC 49239 |
| Flavobacterium psychrophilum JIP02 86 | Gloeobacter violaceus PCC 7421 | Haloterrigena turkmenica DSM 5511 |
| Flexistipes sinusarabici DSM 4947 | Gluconacetobacter diazotrophicus PAI 5 | Halothermothrix orenii H 168 |
| Fluviicola taffensis DSM 16823 | Gluconacetobacter xylinus NBRC 3288 | Halothiobacillus neapolitanus c2 |
| Francisella novicida U112 | Gluconobacter oxydans 621H | Helicobacter acinonychis str Sheeba |
| Francisella philomiragia subsp philomiragia ATCC 25017 | Gordonia bronchialis DSM 43247 | Helicobacter bizzozeronii CIII 1 |
| Francisella sp TX077308 | Gramella forsetii KT0803 | Helicobacter felis ATCC 49179 |
| | Granulibacter bethesdensis CGDNIH1 | Helicobacter hepaticus ATCC 51449 |
| | | Helicobacter mustelae 12198 |
| | | Helicobacter pylori 26695 |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---|
| Heliobacterium modesticaldum Ice1 | Kytococcus sedentarius DSM 20547 | Legionella pneumophila subsp pneumophila ATCC 43290 |
| Herbaspirillum seropedicae SmR1 | Lacinutrix sp 5H 3 7 4 | Leifsonia xyli subsp xyli str CTCB07 |
| Hermiimonas arsenicoydans | Lactobacillus acidophilus NCFM | Leptospira biflexa serovar Patoc strain Patoc 1 Ames |
| Herpetosiphon aurantiacus DSM 785 | Lactobacillus amylovorus GRL 1112 | Leptospira borgpetersenii serovar Hardjo bovis str L550 |
| Hippea maritima DSM 10411 | Lactobacillus buchneri NRRL B 30929 | Leptospira interrogans serovar Copenhageni str Fiocruz L1 130 |
| Hirschia baltica ATCC 49814 | Lactobacillus casei ATCC 334 | Leptothrix cholodnii SP 6 |
| Hydrogenobacter thermophilus TK 6 | Lactobacillus crispatus ST1 | Leptotrichia buccalis C 1013 b |
| Hydrogenobaculum sp Y04AAS1 | Lactobacillus delbrueckii subsp bulgaricus ATCC 11842 | Leuconostoc citreum KM20 |
| Hyperthermus butylicus DSM 5456 | Lactobacillus fermentum IFO 3956 | Leuconostoc gasicomitatum LMG 18811 |
| Hyphomicrobium denitrificans ATCC 51888 | Lactobacillus gasseri ATCC 33323 | Leuconostoc kimchii IMSNU 11154 |
| Hyphomicrobium sp | Lactobacillus helveticus DPC 4571 | Leuconostoc mesenteroides subsp mesenteroides ATCC 8293 |
| Hyphomonas neptunium ATCC 15444 | Lactobacillus johnsonii NCC 533 | Leuconostoc sp C2 |
| Idiomarina loihiensis L2TR | Lactobacillus kefiranofaciens ZW3 | Listeria innocua Clip11262 |
| Ignicoccus hospitalis KIN4 I | Lactobacillus plantarum subsp plantarum ST III | Listeria ivanovii |
| Ignisphaera aggregans DSM 17230 | Lactobacillus reuteri DSM 20016 | Listeria monocytogenes EGD e |
| Ilyobacter polytropus DSM 2926 | Lactobacillus rhamnosus ATCC 8530 | Listeria seeligeri serovar 1 2b str SLCC3954 |
| Intrasporangium calvum DSM 43043 | Lactobacillus ruminis ATCC 27782 | Listeria welshimeri serovar 6b str SLCC5334 |
| IsotERICOLA variabilis 225 | Lactobacillus sakei subsp sakei 23K | Lysinibacillus sphaericus C3 41 |
| Isosphaera pallida ATCC 43644 | Lactobacillus salivarius UCC118 | Macrococcus caseolyticus JCSC5402 |
| Jannaschia sp CCS1 | Lactobacillus sanfranciscensis TMW 11304 | Magnetococcus marinus MC 1 |
| Janthinobacterium sp Marseille | Lactococcus garvieae ATCC 49156 | Magnetospirillum magneticum AMB 1 |
| Jonesia denitrificans DSM 20603 | Lactococcus lactis subsp cremoris NZ9000 | Mahella australiensis 50 1 BON |
| Kangiella koreensis DSM 16069 | Laribacter hongkongensis HLHK9 | Mannheimia succiniciproducens MBEL55E |
| Ketogulonicigenium vulgare WSH 001 | Lawsonia intracellularis PHE MN1 00 | Maribacter sp HTCC2170 |
| Kineococcus radiotolerans SRS30216 | Leadbetterella byssophila DSM 17132 | Maricaulis maris MCS10 |
| Kitasatospora setae KM 6054 | Legionella longbeachae NSW150 | Marinithermus hydrothermalis DSM 14884 |
| Klebsiella oxytoca KCTC 1686 | | Marinobacter adhaerens HP15 |
| Klebsiella pneumoniae 342 | | Marinobacter aquaeolei VT8 |
| Klebsiella variicola At 22 | | Marinomonas mediterranea MMB 1 |
| Kocuria rhizophila DC2201 | | |
| Kosmotoga olearia TBF 1951 | | |
| Kribbella flavida DSM 17836 | | |
| Krokinobacter sp 4H 3 7 5 | | |
| Kyrpidia tusciae DSM 2912 | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|---------------------------------------|
| Marinomonas posidonica IVIA Po 181 | Methanohalobium evestigatum Z 7303 | Methylococcus capsulatus str Bath |
| Marinomonas sp MWYL1 | Methanohalophilus mahii DSM 5219 | Methylomicrobium alcaliphilum |
| Marivirga tractuosa DSM 4126 | Methanoplanus petrolearius DSM 11571 | Methylomonas methanica MC09 |
| Megasphaera elsdenii | Methanopyrus kandleri AV19 | Methylotenera mobilis JLW8 |
| Meiothermus ruber DSM 9946 | Methanoregula boonei 6A8 | Methylotenera versatilis 301 |
| Meiothermus silvanus DSM 9946 | Methanosaeta concilii GP6 | Methylovorus glucosetrophus SIP3 4 |
| Melissococcus plutonius ATCC 35311 | Methanosaeta thermophila PT | Methylovorus sp MP688 |
| Mesoplasma florum L1 | Methanosalsum zhilinae DSM 4017 | Micavibrio aeruginosavorus ARL 13 |
| Mesorhizobium ciceri biovar biserrulae WSM1271 | Methanosarcina acetivorans C2A | Microbacterium testaceum StLB037 |
| Mesorhizobium loti MAFF303099 | Methanosarcina barkeri str Fusaro | Micrococcus luteus NCTC 2665 |
| Mesorhizobium opportunistum WSM2075 | Methanosarcina mazei Go1 | Microcystis aeruginosa NIES 843 |
| Metallosphaera cuprina Ar 4 | Methanosphaera stadtmanae DSM 3091 | Microlunatus phosphovorus NM 1 |
| Metallosphaera sedula DSM 5348 | Methanosphaerula palustris E1 9c | Micromonospora aurantiaca ATCC 27029 |
| Methanobacterium sp AL 21 | Methanospirillum hungatei JF 1 | Micromonospora sp L5 |
| Methanobrevibacter ruminantium | Methanothermobacter marburgensis str Marburg | Mobiluncus curtisii ATCC 43063 |
| M1Methanobrevibacter smithii ATCC 35061 | Methanothermobacter thermautotrophicus str Delta H | Moorella thermoacetica ATCC 39073 |
| Methanocaldococcus fervens AG86 | Methanothermococcus okinawensis IH1 | Moraxella catarrhalis RH4 |
| Methanocaldococcus infernus ME | Methanothermus fervidus DSM 2088 | Muricauda ruestringensis DSM 13258 |
| Methanocaldococcus jannaschii DSM 2661 | Methanotorris igneus Kol 5 | Mycobacterium abscessus ATCC 19977 |
| Methanocaldococcus sp FS406 22 | Methylacidiphilum infernorum V4 | Mycobacterium africanum GM041182 |
| Methanocaldococcus vulcanius M7 | Methylibium petroleiphilum PM1 | Mycobacterium avium 104 |
| Methanocella arvoryzae MRE50 | Methylobacillus flagellatus KT | Mycobacterium bovis AF2122 97 |
| Methanocella paludicola SANAE | Methylobacterium chloromethanicum CM4 | Mycobacterium canettii CIPT 140010059 |
| Methanococcoides burtonii DSM 6242 | Methylobacterium extorquens AM1 | Mycobacterium gilvum PYR GCK |
| Methanococcus aeolicus Nankai 3 | Methylobacterium nodulans ORS 2060 | Mycobacterium leprae TN |
| Methanococcus maripaludis S2 | Methylobacterium populi BJ001 | Mycobacterium marinum M |
| Methanococcus vannieli SB | Methylobacterium radiotolerans JCM 2831 | Mycobacterium rhodesiae NBB3 |
| Methanococcus voltae A3 | Methylobacterium sp 4 46 | Mycobacterium smegmatis str MC2 155 |
| Methanocorpusculum labreanum Z | Methylocella silvestris BL2 | Mycobacterium sp JDM601 |
| Methanoculleus marisnigri JR1 | | Mycobacterium tuberculosis H37Rv |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---|--|--|
| Mycobacterium ulcerans Ag99 | Neisseria lactamica 020 06 | Orientia tsutsugamushi str Boryong |
| Mycobacterium vanbaalenii PYR 1 | Neisseria meningitidis M01 240149 | Oscillibacter valericigenes |
| Mycoplasma agalactiae PG2 | Neorickettsia risticii str Illinois | Owenweeksia hongkongensis DSM 17368 |
| Mycoplasma arthritidis 158L3 1 | Neorickettsia sennetsu str Miyayama | Paenibacillus mucilaginosus KNP414 |
| Mycoplasma bovis PG45 | Niastella koreensis GR20 10 | Paenibacillus polymyxa E681 |
| Mycoplasma capricolum subsp capricolum ATCC 27343 | Nitratifactor salsuginis DSM 16511 | Paenibacillus sp JDR 2 |
| Mycoplasma conjunctivae HRC 581 | Nitratiruptor sp SB155 2 | Paenibacillus terrae HPL 003 |
| Mycoplasma crocodyli MP145 | Nitrobacter hamburgensis X14 | Paludibacter propionigenes WB4 |
| Mycoplasma gallisepticum str Rlow | Nitrobacter winogradskyi Nb 255 | Pantoea ananatis LMG 20103 |
| Mycoplasma genitalium G37 | Nitrosococcus halophilus Nc4 | Pantoea sp At 9b |
| Mycoplasma haemocanis str Illinois | Nitrosococcus oceani ATCC 19707 | Pantoea vagans C9 1 |
| Mycoplasma haemofelis str Langford 1 | Nitrosococcus watsonii C 113 | Parabacteroides distasonis ATCC 8503 |
| Mycoplasma hominis | Nitrosomonas europaea ATCC 19718 | Parachlamydia acanthamoebae UV 7 |
| Mycoplasma hyopneumoniae 232 | Nitrosomonas eutropha C91 | Paracoccus denitrificans PD1222 |
| Mycoplasma hyorhinis HUB 1 | Nitrosomonas sp AL212 | Parvibaculum lavamentivorans DS 1 |
| Mycoplasma leachii 99 014 6 | Nitrosopumilus maritimus SCM1 | Parvularcula bermudensis HTCC2503 |
| Mycoplasma mobile 163K | Nitrosospira multiformis ATCC 25196 | Pasteurella multocida subsp multocida str Pm70 |
| Mycoplasma mycoides subsp mycoides SC str PG1 | Nocardia farcinica IFM 10152 | Pectobacterium atrosepticum SCRI1043 |
| Mycoplasma penetrans HF 2 | Nocardioides sp JS614 | Pectobacterium carotovorum subsp carotovorum PC1 |
| Mycoplasma pneumoniae M129 | Nocardiopsis dassonvillei subsp dassonvillei DSM 43111 | Pectobacterium wasabiae WPP163 |
| Mycoplasma pulmonis UAB CTIP | Nostoc azollae 0708 | Pediococcus pentosaceus ATCC 25745 |
| Mycoplasma putrefaciens KS1 | Nostoc punctiforme PCC 73102 | Pedobacter heparinus DSM 2366 |
| Mycoplasma suis str Illinois | Nostoc sp PCC 7120 | Pedobacter saltans DSM 12145 |
| Mycoplasma synoviae 53 | Novosphingobium aromaticivorans DSM 12444 | Pelagibacterium halotolerans B2 |
| Myxococcus fulvus HW 1 | Novosphingobium sp PP1Y | Pelobacter carbinolicus DSM 2380 |
| Myxococcus xanthus DK 1622 | Oceanithermus profundus DSM 14977 | Pelobacter propionicus DSM 2379 |
| Nakamurella multipartita DSM 44233 | Oceanobacillus iheyensis HTE831 | Pelodictyon phaeoclathratiforme BU 1 |
| Nanoarchaeum equitans Kin4 M | Ochrobactrum anthropi ATCC 49188 | Pelotomaculum thermopropionicum SI |
| Natranaerobius thermophilus JW NM WN LF | Odoribacter splanchnicus DSM 20712 | Persephonella marina EX H1 |
| Natrialba magadii ATCC 43099 | Oenococcus oeni PSU 1 | Petrogoga mobilis SJ95 |
| Natronomonas pharaonis DSM 2160 | Oligotropha carboxidovorans OM5 | |
| Nautilia profundicola AmH | Olsenella uli DSM 7084 | |
| Neisseria gonorrhoeae FA 1090 | Onion yellows phytoplasma OY M | |
| | Opitutus terrae PB90 1 | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|--|--|
| Phenylobacterium zucineum HLK1 | Pseudomonas brassicacearum subsp brassicacearum NFM421 | Ramlibacter tataouinensis TB310 |
| Photobacterium profundum SS9 | Pseudomonas entomophila L48 | Renibacterium salmoninarum ATCC 33209 |
| Photorhabdus asymbiotica subsp asymbiotica ATCC 43949 | Pseudomonas fluorescens F113 | Rhizobium etli CFN 42 |
| Photorhabdus luminescens subsp laumondii T101 | Pseudomonas fulva 12 X | Rhizobium leguminosarum bv viciae 3841 |
| Picrophilus torridus DSM 9790 | Pseudomonas mendocina ymp | Rhodobacter capsulatus SB 1003 |
| Pirellula staleyi DSM 6068 | Pseudomonas protegens Pf 5 | Rhodobacter sphaeroides 241 |
| Planctomyces brasiliensis DSM 5305 | Pseudomonas putida F1 | Rhodococcus equi 103S |
| Planctomyces limnophilus DSM 3776 | Pseudomonas stutzeri A1501 | Rhodococcus erythropolis PR4 |
| Polaromonas naphthalenivorans CJ2 | Pseudomonas syringae pv phaseolicola 1448A | Rhodococcus jostii RHA1 |
| Polaromonas sp JS666 | Pseudonocardia | Rhodococcus opacus B4 |
| Polymorphum gilvum SL003B 26A1 | dioxanivorans CB1190 | Rhododerax ferrireducens T118 |
| Polynucleobacter necessarius subsp asymbioticus QLW P1DMWA 1 | Pseudovibrio sp FO BEG1 | Rhodomicrobium vanniellii ATCC 17100 |
| Porphyromonas asaccharolytica DSM 20707 | Pseudoxanthomonas spadix BD a59 | Rhodopirellula baltica SH 1 |
| Porphyromonas gingivalis W83 | Pseudoxanthomonas suwonensis 11 1 | Rhodopseudomonas palustris CGA009 |
| Prevotella denticola F0289 | Psychrobacter arcticus 273 4 | Rhodospirillum centenum SW |
| Prevotella melaninogenica ATCC 25845 | Psychrobacter cryohalolentis K5 | Rhodospirillum rubrum ATCC 11170 |
| Prevotella ruminicola 23 | Psychrobacter sp PRwf 1 | Rhodothermus marinus DSM 4252 |
| Prochlorococcus marinus str MIT 9215 | Psychromonas ingrahamii 37 | Rickettsia africae ESF 5 |
| Propionibacterium acnes 6609 | Pusillimonas sp T7 7 | Rickettsia akari str Hartford |
| Propionibacterium freudenreichii subsp shermanii CIRM BIA1 | Pyrobaculum aerophilum str IM2 | Rickettsia bellii RML369 C |
| Prosthecochloris aestuarii DSM 271 | Pyrobaculum arsenaticum DSM 13514 | Rickettsia canadensis str McKiel |
| Proteus mirabilis HI4320 | Pyrobaculum calidifontis JCM 11548 | Rickettsia conorii str Malish 7 |
| Pseudoalteromonas atlantica T6c | Pyrobaculum islandicum DSM 4184 | Rickettsia felis URRWXC2 |
| Pseudoalteromonas haloplanktis TAC125 | Pyrobaculum neutrophilum V24Sta | Rickettsia heilongjiangensis 054 |
| Pseudoalteromonas sp SM9913 | Pyrobaculum sp 1860 | Rickettsia japonica YH |
| Pseudogulbenkiania sp NH8B | Pyrococcus abyssi GE5 | Rickettsia massiliae MTU5 |
| Pseudomonas aeruginosa PAO1 | Pyrococcus furiosus DSM 3638 | Rickettsia peacockii str Rustic |
| | Pyrococcus horikoshii OT3 | Rickettsia prowazekii str Madrid E |
| | Pyrococcus sp NA2 | Rickettsia rickettsii str Iowa |
| | Pyrococcus yayanosii CH1 | Rickettsia sibirica 246 |
| | Pyrolobus fumarii 1A | Rickettsia slovacica 13 B |
| | Rahnella sp Y9602 | Rickettsia typhi str Wilmington |
| | Ralstonia eutropha JMP134 | Riemerella anatipestifer ATCC 11845 DSM 15868 |
| | Ralstonia pickettii 12D | Robiginitalea biformata HTCC2501 |
| | Ralstonia solanacearum GMI1000 | Roseburia hominis A2 183 |
| | | Roseiflexus castenholzii DSM 13941 |
| | | Roseiflexus sp RS 1 |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|--|---|---|
| Roseobacter denitrificans OCh 114 | Shewanella piezotolerans WP3 | Stackebrandtia nassauensis DSM 44728 |
| Roseobacter litoralis Och 149 | Shewanella putrefaciens CN 32 | Staphylococcus aureus subsp aureus JH1 |
| Rothia dentocariosa ATCC 17931 | Shewanella sediminis HAW EB3 | Staphylococcus carnosus subsp carnosus TM300 |
| Rothia mucilaginosa DY 18 | Shewanella sp ANA 3 | Staphylococcus epidermidis ATCC 12228 |
| Rubrobacter xylanophilus DSM 9941 | Shewanella violacea DSS12 | Staphylococcus haemolyticus JCSC1435 |
| Ruegeria pomeroyi DSS 3 | Shewanella woodyi ATCC 51908 | Staphylococcus lugdunensis HKU09 01 |
| Ruegeria sp TM1040 | Shigella boydii Sb227 | Staphylococcus pseudintermedius HKU10 03 |
| Ruminococcus albus 7 | Shigella dysenteriae Sd197 | Staphylothermus hellenicus DSM 12710 |
| Runella slithyformis DSM 19594 | Shigella flexneri 2a str 301 | Staphylothermus marinus F1 |
| Saccharomonospora viridis DSM 43017 | Shigella sonnei Ss046 | Starkeya novella DSM 506 |
| Saccharophagus degradans 2 40 | Sideroxydans lithotrophicus ES 1 | Stenotrophomonas maltophilia K279a |
| Saccharopolyspora erythraea NRRL 2338 | Simkania negevensis Z | Stigmatella aurantiaca DW4 3 1 |
| Salinibacter ruber DSM 13855 | Sinorhizobium fredii NGR234 | Streptobacillus moniliformis DSM 12112 |
| Salinispora arenicola CNS 205 | Sinorhizobium medicae WSM419 | Streptococcus agalactiae 2603V R |
| Salinispora tropica CNB 440 | Sinorhizobium meliloti 1021 | Streptococcus dysgalactiae subsp equisimilis ATCC 12394 |
| Salmonella bongori NCTC 12419 | Slackia heliotrinireducens DSM 20476 | Streptococcus equi subspecies zooepidemicus |
| Salmonella enterica subsp enterica serovar Enteritidis str P125109 | Sodalis glossinidius str morsitans | Streptococcus gallolyticus subsp gallolyticus ATCC BAA 2069 |
| Sanguibacter keddiei DSM 10542 | Sorangium cellulosum So ce56 | Streptococcus gordonii str Challis substr CH1 |
| Sebaldeia termitidis ATCC 33386 | Sphaerobacter thermophilus DSM 20745 | Streptococcus macedonicus ACA DC 198 |
| Segniliparus rotundus DSM 44985 | Sphaerochaeta coccoides DSM 17374 | Streptococcus mitis B6 |
| Selenomonas sputigena ATCC 35185 | Sphaerochaeta globus str Buddy | Streptococcus mutans NN2025 |
| Serratia plymuthica AS9 | Sphaerochaeta pleomorpha str Grapes | Streptococcus oralis Uo5 |
| Serratia proteamaculans 568 | Sphingobacterium sp 21 | Streptococcus parasanguinis ATCC 15912 |
| Serratia sp AS12 | Sphingobium chlorophenolicum L 1 | Streptococcus parauberis KCTC 11537 |
| Serratia symbiotica str Cinara cedri | Sphingobium japonicum UT26S | Streptococcus pasteurianus ATCC 43144 |
| Shewanella amazonensis SB2B | Sphingobium sp SYK 6 | Streptococcus pneumoniae ST556 |
| Shewanella baltica OS155 | Sphingomonas wittichii RW1 | Streptococcus pseudopneumoniae IS7493 |
| Shewanella denitrificans OS217 | Sphingopyxis alaskensis RB2256 | |
| Shewanella frigidimarina NCIMB 400 | Spirochaeta caldaria DSM 7334 | |
| Shewanella halifaxensis HAW EB4 | Spirochaeta smaragdinae DSM 11293 | |
| Shewanella loihica PV 4 | Spirochaeta thermophila DSM 6192 | |
| Shewanella oneidensis MR 1 | Spirosoma linguale DSM 74 | |
| Shewanella pealeana ATCC 700345 | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | | |
|---|---|--|
| Streptococcus pyogenes M1 GAS | Synechocystis sp PCC 6803 | Thermococcus gammatolerans EJ3 |
| Streptococcus salivarius JIM8780 | Syntrophobacter fumaroxidans MPOB | Thermococcus kodakarensis KOD1 |
| Streptococcus sanguinis SK36 | Syntrophobotulus glycolicus DSM 8271 | Thermococcus onnurineus NA1 |
| Streptococcus suis 05ZYH33 | Syntrophomonas wolfei subsp wolfei str Goettingen | Thermococcus sibiricus MM 739 |
| Streptococcus thermophilus CNRZ1066 | Syntrophothermus lipocalidus DSM 12680 | Thermococcus sp 4557 |
| Streptococcus uberis 0140J | Syntrophus aciditrophicus SB | Thermocrinis albus DSM 14484 |
| Streptomyces avermitilis MA 4680 | Tannerella forsythia ATCC 43037 | Thermodesulfatator indicus DSM 15286 |
| Streptomyces bingchenggensis BCW 1 | Taylorella equigenitalis MCE9 | Thermodesulfobacterium sp OPB45 |
| Streptomyces cattleya NRRL 8057 DSM 46488 | Tepidanaerobacter acetatoxydans Re1 | Thermodesulfobium narugense DSM 14796 |
| Streptomyces coelicolor A32 | Teredinibacter turnerae T7901 | Thermodesulfobivrio yellowstonii DSM 11347 |
| Streptomyces flavogriseus ATCC 33331 | Terriglobus saanensis SP1PR4 | Thermofilum pendens Hrk 5 |
| Streptomyces griseus subsp griseus NBRC 13350 | Tetragenococcus halophilus Thauera sp MZ1T | Thermomicrobium roseum DSM 5159 |
| Streptomyces scabiei 8722 | Thermaerobacter marianensis DSM 12885 | Thermomonospora curvata DSM 43183 |
| Streptomyces sp SirexAA E | Thermanaerovibrio acidaminovorans DSM 6589 | Thermoplasma acidophilum DSM 1728 |
| Streptomyces violaceusniger Tu 4113 | Thermincola potens JR | Thermoplasma volcanium GSS1 |
| Streptosporangium roseum DSM 43021 | Thermoanaerobacter brockii subsp finii Ako 1 | Thermoproteus uzoniensis 768 20 |
| Sulfobacillus acidophilus DSM 10332 | Thermoanaerobacter italicus Ab9 | Thermosediminibacter oceani DSM 16646 |
| Sulfolobus acidocaldarius DSM 639 | Thermoanaerobacterium thermosaccharolyticum DSM 571 | Thermosiphon africanus TCF52B |
| Sulfolobus islandicus M1425 | Thermoanaerobacterium xylanolyticum LX 11 | Thermosiphon melanesiensis BI429 |
| Sulfolobus solfataricus P2 | Thermoanaerobacter mathranii subsp mathranii str A3 | Thermosphaera aggregans DSM 11486 |
| Sulfolobus tokodaii str 7 | Thermoanaerobacter pseudethanolicus ATCC 33223 | Thermosynechococcus elongatus BP 1 |
| Sulfuricurvum kujiense DSM 16994 | Thermoanaerobacter sp X513 | Thermotoga lettingae TMO |
| Sulfurihydrogenibium azorense Az Fu1 | Thermoanaerobacter tengcongensis MB4 | Thermotoga maritima MSB8 |
| Sulfurihydrogenibium sp YO3AOP1 | Thermoanaerobacter wiegelii R18B1 | Thermotoga naphthophila RKU 10 |
| Sulfurimonas autotrophica DSM 16294 | Thermobaculum terrenum ATCC BAA 798 | Thermotoga neapolitana DSM 4359 |
| Sulfurimonas denitrificans DSM 1251 | Thermobifida fusca YX | Thermotoga petrophila RKU 1 |
| Sulfurospirillum deleyianum DSM 6946 | Thermobispora bispora DSM 43833 | Thermotoga sp RQ2 |
| Sulfurovum sp NBC37 1 | Thermococcus barophilus MP | Thermotoga thermarum DSM 5069 |
| Symbiobacterium thermophilum IAM 14863 | | Thermovibrio ammonificans HB 1 |
| Synechococcus elongatus PCC 6301 | | Thermovirga lienii DSM 17291 |
| Synechococcus sp CC9605 | | |

SIANN: Strain Identification by Alignment to Near Neighbors

| | |
|-----------------------------------|--------------------------------|
| Thermus scotoductus SA 01 | Vibrio parahaemolyticus RIMD |
| Thermus sp CCB US3 UF1 | 2210633 |
| Thermus thermophilus HB27 | Vibrio sp EJY3 |
| Thioalkalimicrobium cyclicum | Vibrio splendidus LGP32 |
| ALM1 | Vibrio vulnificus CMCP6 |
| Thioalkalivibrio sp K90mix | Vulcanisaeta distributa DSM |
| Thioalkalivibrio sulfidophilus HL | 14429 |
| EbGr7 | Vulcanisaeta moutnovskia |
| Thiobacillus denitrificans | 768 28 |
| ATCC 25259 | Waddlia chondrophila WSU |
| Thiomicrospira crunigena | 86 1044 |
| XCL 2 | Weeksella virosa DSM 16922 |
| Thiomonas intermedia K12 | Weissella koreensis KACC |
| Tolomonas auensis DSM 9187 | 15510 |
| Treponema azotonutricium | Wigglesworthia glossinidia |
| ZAS 9 | endosymbiont of Glossina |
| Treponema brennaboreense | brevipalpis |
| DSM 12168 | Wolbachia endosymbiont of |
| Treponema denticola ATCC | Culex quinquefasciatus Pel |
| 35405 | Wolbachia sp wRi |
| Treponema pallidum subsp | Wolinella succinogenes DSM |
| pallidum str Nichols | 1740 |
| Treponema paraluis-cuniculi | Xanthobacter autotrophicus |
| Cuniculi A | Py2 |
| Treponema primitia ZAS 2 | Xanthomonas albilineans GPE |
| Treponema succinifaciens | PC73 |
| DSM 2489 | Xanthomonas axonopodis pv |
| Trichodesmium erythraeum | citri str 306 |
| IMS101 | Xanthomonas campestris pv |
| Tropheryma whipplei str Twist | campestris str 8004 |
| Truepera radiovictrix DSM | Xanthomonas oryzae pv |
| 17093 | oryzae KACC 10331 |
| Tsukamurella paurometabola | Xenorhabdus bovienii SS 2004 |
| DSM 20162 | Xenorhabdus nematophila |
| Ureaplasma parvum serovar 3 | ATCC 19061 |
| str ATCC 27815 | Xylanimonas cellulolytica |
| Ureaplasma urealyticum | DSM 15894 |
| serovar 10 str ATCC 33699 | Xylella fastidiosa 9a5c |
| Variovorax paradoxus S110 | Yersinia enterocolitica subsp |
| Veillonella parvula DSM 2008 | enterocolitica 8081 |
| Verminephrobacter eiseniae | Yersinia pestis CO92 |
| EF01 2 | Yersinia pseudotuberculosis IP |
| Verrucospora maris AB 18 | 32953 |
| 032 | Zobellia galactanivorans |
| Vibrio anguillarum 775 | Zunongwangia profunda SM |
| Vibrio cholerae O1 biovar El | A87 |
| Tor str N16961 | Zymomonas mobilis subsp |
| Vibrio fischeri ES114 | mobilis ATCC 10988 |
| Vibrio furnissii NCTC 11218 | |
| Vibrio harveyi ATCC BAA 1116 | |